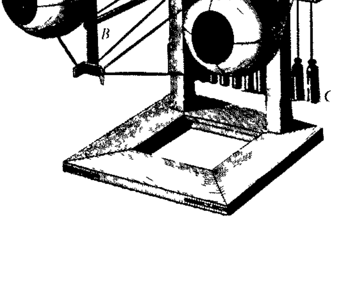
PRIMARY PERCEPTION OF STIMULUS STRUCTURE

LUCAS MENS



Primary Perception of Stimulus Structure

Primary Perception of Stimulus Structure

Een wetenschappelijke proeve
op het gebied van de Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit te Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op
maandag 31 october 1988 des namiddags te 1.30 uur precies

door:

Lucas Hermanus Maria Mens

geboren op 20 december 1956 te Vught

N I C I
Nijmeegs Instituut voor Cognitie-onderzoek
en Informatietechnologie

Promotor: Prof.dr. A.J.W.M. Thomassen
Co-referent: Dr. E.L.J. Leeuwenberg

Dankbetuiging

Bij de meeste hoofdstukken treft U in een voetnoot enige namen aan van personen die bijdroegen aan het betreffende werk. Moge die voetnoot telkens aandacht van de lezer krijgen, want het is een te nederige plaats om te bedanken voor een vaak essentiële inbreng.

Dit is een proefschrift uit de school van de Nijmeegse vakgroep Functieleer. Ik wil bevestigen dat aan deze school bijzonder goede leraren verbonden zijn. In hen verenigt zich vakkundigheid en persoonlijkheid. Voorzover ik in staat was hiervan te profiteren dank ik hen; mijn schooltijd is in ieder geval immer boeiend geweest.

Zeer erkentelijk ben ik mijn klasgenoten van het uit ZWO fondsen betaalde project Structurele Informatie Theorie. Vooral met Peter van der Helm heb ik veel stof kunnen doornemen. Bovendien is het vervaardigen van dit boek door hem zeer vergemakkelijkt.

Contents

Introduction

Theoretical accounts of stimulus structure can be used to predict the perceptual organization of stimuli. This thesis examines the perceived organization of simple line drawings using Leeuwenberg's (1971) Structural Information Theory (SIT) as a starting point. With respect to the recognition of spoken words, the importance of the rhythmic organization of sentences as predicted by Martin (1972) is investigated.

Three issues will be considered in this thesis. The first question is whether an experimental method can be developed to decide between two options: the first option is that the perceived organization is **inherent** to the perceptual representation (or the perceptual code) of stimulus structure. The alternative option is that the perceived organization is one that is added in later stages (e.g., while reflecting on one's impressions). In the latter case, organization as perceived can be called an epiphenomenon; the actual perceptual process could involve a completely different organization, or even none at all (c.f. classification on the basis of distinctive features or template matching). To decide between these options in our experiments, we will use SIT to predict effects of organization in *primary perception of structure* (the initial access to the representation of structure), while separating such effects from secondary organizational effects due to prolonged or *post-access* processing. Note that "access" is used in this thesis in the sense of "becoming available", rather than denoting the activation of an already established representation. If the organizational aspect can be shown to belong to primary perception of structure, a second and a third issue becomes relevant.

The second issue is concerned with organization and the **completeness** of representations. Most patterns, even very simple ones, allow alternative perceptual organizations. On page 7 it is shown that alternative organizations of a pattern correspond to alternative descriptions in terms of a coding language such as SIT, and that each of these descriptions is incomplete in the sense that it does not allow a reconstruction of the pattern including all structural regularities that were initially present. To the extent that SIT codes reflect perceptual representations, this could imply that the visual system has to choose between alternative perceptual representations of such a pattern, none of which fully specifies its structure. However, this theoretical incompleteness of each separate representation does not fit in as such with our ability to make an accurate drawing of (simple) patterns after viewing. That is, it is in conflict with the fact that the structure of the particular pattern seems to be represented uniquely and completely. This conflict would be resolved by assuming the existence of a second representation apart from the one that is phenomenally dominant, supplying all the information that is lacking in the dominant one. In Structural Information Theory, this second representation is called *complementary*. The second empirical question is: can it be shown that the complementary representation is concurrently present in primary perception although it is suppressed by the dominant one?

This brings us to the third issue. If it can be shown that the internal structure of the stimulus is completely represented in primary perception, it becomes interesting to ask whether the relation between its proper structure and that of earlier perceived patterns (its *external structure*) is equally important. In other words, the third question is whether the perceived organization depends entirely upon an analysis of the stimulus structure that is immediately given, or instead is subject to substantial change due to context. This is called the issue of the **autonomy** of perception.

In the following sections of this introduction, theoretical notions underlying the three central issues will be considered. It will be argued that perception must be equivalent to making a description of regularities in the environment, rather than an analogue registration of geometrical information. It will subsequently be shown that the notion of organization is essential to understand how perceivers arrive at a perceptual representation of visual and auditory structure. Structural Information Theory will be introduced as an advanced attempt to quantify stimulus structure and a convenient means to demonstrate and define completeness and the principle of complementarity. Subsequently, previous attempts to test the first two of the above questions (inherence and completeness) will be discussed, emphasizing an experimental method that yields effects of primary perception and a form of data analysis that corrects for post-access biases. This particular method and analysis will turn out to be particularly useful to investigate the third question, the autonomy of the perception of stimulus structure. Finally, the specific question addressed in the five following chapters is stipulated as well as the respective experimental approach to be followed in each of them.

Organization

The production of our perception is an organized world. On a large scale, we consciously distinguish objects and establish meaningful relations between these objects. Thus, mental representations at the highest "cognitive" level can be understood as structural descriptions of the environment (Minsky, 1975). The power of this idea is manifest in working models of parts of human cognition: the expert systems of artificial intelligence (Minsky, 1975).

However, to what extent is it also productive to regard *perceptual* representations as being equivalent to structural descriptions? We are generally less conscious of an internal structure within the perceived objects; as a rule, we experience a particular object or event as a unified whole. Accordingly, the process of perception, following which we arrive at the mental representations, has been seen as the formation of a pattern of brain activity which is somehow isomorphic to the object. This is what Gestalt psychology has suggested (Koffka, 1935) and what reappears in the distributed representation in some (neural) network models (Ballard, Hinton & Sejnowski, 1983). Distributed representations are naturally coupled to distributed *processing*. As such, these models indeed have attractive properties which line up with work in computational vision research (Marr, 1982) on the processing of local aspects of the image (e.g., surface properties such as distance and slant). At present, however, the notion of distributed representations does not seem to advance our understanding of clearly *non-local* perceptual aspects such as organization, beyond what has been formalized in feature-recognition models (Pinker, 1984). Specifically, it appears that for any local feature (for example a T-junction) that can be specified, a pattern can be found in which that feature is physically present but perceptually hidden by an (atypical) embedding context (cf. Buffart, Leeuwenberg & Restle, 1981); these models are hard pressed to account for these global, organizational effects in a consistent way.

Alternatively, perception may be compared to the production of a *structural description*: an expression specifying parts and (spatial) relationships among them. This seems to be the appropriate approach to model what must be seen as the central task of perception: the classification of a pattern as an instance of a set of structurally equivalent ones. For instance, in reading we do not respond to a particular retinal stimulation, but to an equivalence class of stimulations which relate to a specific perceptual category (e.g., a particular character) in a many-to-one fashion (Neisser, 1967). Sutherland (1968) listed a number of feats of the perceptual system supporting the description metaphor. These feats are the abstraction from variations in brightness, size, retinal position, local distortions, type of contour and surface. This becomes evident, for example, in our stable perception of the global organization of complex scenes under variation of the components, in confusions of structurally similar shapes, and in perceptual learning via schemata. The importance of the structural description metaphor is also testified by the effect of different reference frames (Rock, 1983; Palmer, 1983) and by the capacity to segment a picture in different ways. An important class of theories based on structural descriptions are *coding models.*

Coding models

The explicit description of structure is the aim of coding models. Coding models were first developed within limited domains. For instance, Simon and Kotovsky (1963) quantified the complexity of Thurstone letter series used in intelligence tests. Series were coded by capturing internal regularities using a small set of coding rules. The complexity of the resulting code was used as a predictor of the complexity of the series. Combining neurophysiological findings and early work on picture languages used in machine vision, Sutherland (1968) was one of the first to sketch a coding model of non-linear (two and three dimensional) visual patterns. He made the essential step from the image to a code by writing a simple line drawing as a linear sequence of symbols. These symbols stand for the primitive components of the pattern, in this case lines and angles, defined with Hubel and Wiesel's feature detectors in mind.

In principle, a symbol sequence may preserve an arbitrary amount of metrical information. A symbol may stand for a particular line of an exact length under a given angle at a specific place in the visual field. However, it is a basic tenet for coding models (McKay, 1950; Leeuwenberg, 1973) that describing the structure of a pattern presupposes the abstraction from particular instances in two ways. In the first place, different lines of about—for example—the same length have to be assigned the same symbol. Secondly, if two components are not taken to be identical, different symbols have to be assigned to these two components, merely indicating their non-equivalence without preserving how much they differ. In other words, structural descriptions in these models are based on the information that the components are either identical or different: they are based on identity relations (Buffart, Leeuwenberg & Restle, 1983).

But we do not have a description, or *code*, without coding rules. Ignoring metrical differences, the identity relations implicit in the pattern are actually described by the application of coding rules. For example, the pattern "A A A B B" can obviously be described—given the two "abstraction conditions" listed above—by the code "3 *(A),2 *(B)" (assuming iteration as a coding rule). This implies an organization of the pattern which incorporates "A A A" as one part and "B B" as another one. Although the number of possible coding rules is vast—and coding models indeed differ in their choice of rules—it is clear that the application of coding rules results in a description which defines a specific organization of the patterns. Thus, organization seems to be a general consequence of expressing identity relations in structural descriptions, not specific to models which map a visual pattern onto a linear sequence of symbols, and independent of the set of coding rules actually endorsed.

As Sutherland did, SIT starts with a translation of the pattern into a linear sequence of descriptive symbols (standing for angles and lines). Coding rules can then be understood as *reduction rules*, reducing the initial sequence to a shorter one with fewer descriptive symbols. Note that this is indeed a technical encoding-scheme, not a model of a two-phased process of perception. The result of this reduction phase is the maximally reduced or *final* code upon which to base experimental predictions, e.g., about the preferred interpretation of patterns. To do so, the reduction rules have to be applied in an exhaustive fashion. In contrast to models emphasizing economy of the encoding process (see Hatfield & Epstein, 1985; see also chapter 2), SIT does not assume any fixed order of application of the reduction rules. Different exhaustive paths of encoding lead to different final codes. The attractive feature of having more than one code of a pattern is that it fits nicely into the ability of (visual) perception to organize most (line-) patterns in several different ways. The problem is that, in general, we perceive only a single organization with a strong preference over all other ones. Thus, a criterion is needed to select which final code of all possible ones is the *best*. In other words, the question to be answered is: what is the guiding principle of the encoding mechanism?

Minimum principle

A central assumption of Leeuwenberg's coding theory, SIT, and one alluded to by Sutherland, is that the *best* code is the one with the smallest number of descriptive elements: the *minimum code*. This selection criterion is the application of a very general principle: the *minimum principle*, which states that the human visual system is directed towards the simplest representation possible of a given pattern. The minimum principle is rooted in the notion of *Prægnanz* underlying the various Gestalt Laws of perceptual organization. The principle was invoked by Hochberg and McAllister (1953) who gave it its present name. These authors sought to predict the perceived two or three-dimensionality of drawings of Necker cubes and similar patterns. The minimum principle as propagated by coding theories is much more than a technical prescription of how to derive codes. Least of all, it is the prediction that perception favours simple representations as such. Instead, it is closely linked to what we stated to be the fundamental condition for representing stimulus structure: the specification of identity relations between descriptive symbols referring to pattern elements. In terms of an encoding scheme, many identity relations are covered in the initial symbol sequence, but only implicitly. In contrast, in the best, or minimum-code, a maximal amount of identity relations is expressed explicitly. In other words, the minimum code of a pattern describes more identity relations than another one, which will always contain more descriptive elements. Collard and Buffart (1983) have given a formal exposition of this idea. In their work, an important consequence of code reduction becomes clear: the incompleteness of a code and hence the need for a *complementary* code.

Complementary organization

Consider, for example, the pattern "A A B A". It contains three identity relations between the three pairs of A's. Assume the following set of (two) rules: iteration and symmetry (s). Then three codes are possible. One code expresses the iterated appearance of the first two A's: "2 *(A) X Y"; each of the other two expresses a symmetry: a symmetry of two A's around one different element in "X s(A,Y)" and a symmetry of two A's around two different elements in "s(A,(XY))". By now it will be clear why the elements not covered by a rule are replaced by "X" and "Y": the only feature covered of these elements is that they are not identical to any of the other elements. The important point is that each of these codes corresponds to a different, mutually exclusive, organization of the symbol sequence. According to SIT, each of these codes is incomplete. In the iteration code, for example, the information that the last element is identical to the first two elements is absent. As far as this code is concerned, "A A B A" cannot be distinguished from "A A B C". However, perceivers are able to discriminate and retain such patterns, and much more complex ones at that. In principle, for each particular pattern a (complex) reduction rule can be defined which completely covers all of the structural regularities in the pattern. Thus, there is nothing that forbids one to invent a coding rule "xxyx" in order to describe pattern "A A B C" completely by: "xxyx(A,B)". However, in that case a very large set of reduction rules has to be assumed, in fact, one for each class of patterns. Instead of one 2-D or 3-D template for each perceptual object (Neisser, 1967), one would have one rule. This is of course implausible. If, alternatively, only a limited number of rules is allowed (see Buffart, 1987, and Van der Helm & Leeuwenberg, 1988, for a formal and psychological foundation of the choice of the limited set of rules used in Structural Information Theory) the incompleteness of many [1] structural codes has to be solved, or compensated for, in another way. Such incompleteness would be compensated for if, besides the minimum code, other codes are supplied, covering all of the structural identities not covered by each one of them alone. Collard and Buffart (1983) emphasize the need for an extra code and named it the principle of *complementarity*. Preferring the latter account of the apparent

[1] SIT, for example, claims that the codes of about 90% of all patterns with 10 descriptive elements are incomplete (Mellink & Buffart, 1987).

completeness of perception, our conclusion is that if perception is engaged in the specification of regularity based on identity relations, more than one code, or representation, should be entertained for the majority of patterns.

Primary perception of structure

For a considerable number of perceptual tasks, successful predictions have been made on the basis of the minimum code as to which organization human observers prefer to assign to a given stimulus (for an overview, see Leeuwenberg & Buffart 1983). Furthermore, the frequency with which the preferred and the second-best organization are reported could be related to the relative complexity of the corresponding codes, which supports the notion of complementarity. But what is the relevance of these successes for our understanding of human perception?

The subjects in most of these studies indicated which organization they preferred after prolonged viewing. It may be suspected that part of their responses reflect an ad-hoc strategy to accommodate to the inferred task-demands ("be geometrically logical", "be creative" etc.); the subject is given ample time to go through a complex, piecemeal confirmation procedure of these "cognitive" solutions. In other words, in order to gain insight in the actual perceptual representation, it is necessary to employ an experimental method which allows one to detect and eliminate such ad-hoc, post-access strategies.

"Cognitive" solutions are less likely to confound the results in studies following the Gottschaldt tradition by using variations of the embedded-figures test (Reed & Johnsen, 1975; Bower & Glass, 1976; Palmer, 1977; Van Tuijl, 1980). Typically, a line drawing is shown that allows multiple organizations: the target pattern. This is followed by a test pattern of which subjects are asked to decide as fast or as accurately as possible whether or not it is a component of the target pattern. The main point of chapter 4 is to show that such a task still yields a multitude of effects, some of which are clearly unrelated to the perceived organization of the target pattern. For example, test patterns with a "good" shape may yield a bias to respond "yes"; similarity of local details of target and test pattern may have the same effect.

The possible confoundings due to strategies and biases are even more important if one wants to decide whether the postulated complementary codes correspond to simultaneously present representations, or to subsequent representations, or to no distinct psychological entities at all. It is clear that this issue cannot be settled by the relative frequency of responses accumulated over many trials, patterns and subjects. A more stringent test seems to be required.

SIT has often been advocated as a representation theory: a theory about the final result of perception. How this result is reached is not the domain of SIT. Buffart (1987) and Restle (1982) take this argument further and state that one should not tap responses that depend only upon the first few seconds of processing: these are deemed to reflect "early" perception governed by non-structural properties of (sensory) mechanisms. Considerable time might be needed to reach the true minimum interpretation. Yet, we give subjects only a minimum amount of time. By doing so, we want to look at a hypothetical stage, which we will call *primary perception of structure*: the first moment that the representation of stimulus structure becomes available, or, as it is called in this thesis, the initial access to that representation. Therefore, we want to exclude all effects of semantic, associative or other memory factors operating in "post-access" stages. The principal method to study this stage has two elements: (a) a very fast presentation of two patterns (for example, within 50 milliseconds) probing early representations and (b) a correction procedure eliminating effects of prolonged processing resulting in preferences for particular perceptual alternatives and in response biases.

To demonstrate the existence of *primary perception of structure* is the main goal of this thesis. It is, however, altogether possible that such a stage of representation does not exist in a recognizable form, distinguishable from the rest of the system. For instance, Marr's (1982) "primal sketch" is the highest level of description in his model that can be recovered from the image only, but it describes local surface information rather than organized structural components. Witkin and Tenenbaum (1983, p. 532), on the other hand, argue in favour of a truly perceptual representation that does instantiate an organization. They contrast the primal sketch stage with "intermediate" vision: the structural decomposition transcending the representation of local surface characteristics, but still without the need to draw on specialized, semantic knowledge.

The term *primary perception* suggests that the relevant processes are finished within a fixed time, say within the 150 milliseconds generally thought necessary for a conscious impression to arise, after which the initial organization may not be found in an unconfounded status. There are two reasons why this may not be so. First, the primary structural representation need not be erased by later interpretations, but, as a rule, may instead be incorporated into more elaborated representations (see chapter 4). Second, it depends on the particular display—for instance, on the absence of noise, on the complexity etc.— whether or not longer periods of scanning a pattern are necessary to "sample" enough information for any organization to arise. However, we will focus on early processing effects in a specific experimental paradigm (using a very short presentation and very simple patterns that can be perceived "at a single glance" in chapters 1, 2 and 4) in order to avoid the above-mentioned problems connected with less restricted paradigms. On the response side, effects of prolonged processing are eliminated by applying the response-bias correction that will be developed in chapter 2 and extended in chapter 4. Unbiased reaction times to "same" responses are measured in chapter 3. In the case of speech perception (chapter 5), reaction times are measured in a phoneme-monitoring task, which is a so-called *on-line* technique, adopted also with the objective of increasing the sensitivity to early processing effects.

We will now present a brief overview of the goal and method of each of the studies described in chapters 1 to 5.

In Part I, consisting of chapter 1, it is attempted to link early processing effects to perceived organization in order to demonstrate that organization is inherent to perception rather than an epiphenomenon. As in the embedded figures test, a complete pattern is presented followed by one out of several possible subpatterns. At the end of such a trial, the subject is shown all subpatterns and then is required to indicate which one of the subpatterns has actually been presented. The difference from the traditional embedded figures test is that recognition is measured at about threshold level: the two patterns are each presented for only 10 milliseconds and they are separated by a very short interval. This is equivalent to a masking paradigm. Accordingly, short intervals yield poor scores, whereas with long intervals the task is much easier. On the basis of Structural Information Theory pairs of a complete pattern and a subpattern are selected which, presented in one order, have a mutually exclusive organization. Presented in the reversed order, however, the perceptual representation of the first pattern is predicted to be able to reinforce a latent interpretation of the second one. By systematic variation, the shortest interval can be found at which a comparison between the two orders of presentation results in an effect on the correct recognition of the subpattern. This difference can be assigned to the role of pattern organization. Thus, an impression is gained of the time needed to process the stimuli from a sensory registration into a representation that implies a perceptual organization.

Completeness is the issue addressed in Part II, comprising chapters 2 and 3. In chapter 2, a complete pattern and a subpattern are presented, like in chapter 1. The purpose now is to investigate the presence of the non-preferred second-best code which, according to the principle of complementarity, has the special function of eliminating the incompleteness of the preferred code. If it has that function, we expect the second-best organization to correspond to a suppressed but *concurrently present* interpretation, instead of merely reflecting a *subsequent* alternative interpretation which happens to be preferred less often. Others have endorsed the notion of subsequent interpretations, even for ambiguous or "multistable" patterns, such as the Necker cube with its reversible depth organization. Clearly, if a non-preferred organization can be shown to be concurrently present, the alternative organizations of a perfectly ambiguous pattern may be assumed to be concurrently present too. In order to exclude confounding effects not belonging to primary perception, a bias-correction procedure based on the distribution of the errors is introduced.

According to SIT, the complementary code considered in chapter 2 contains the necessary information to achieve the complete description of the structure of a pattern. Just like the minimum code, it does not contain *metrical* information. A special type of complementary code is considered for the simple linear arrays of identical elements that will be used in chapter 3. In order to discriminate between such arrays, metrical information (the number of elements) *is* essential. Consider an array of regularly spaced identical characters, for instance two rows of letters, one of three A's and one of five. It is informative that the two arrays have an identical number of elements, but structurally, the size of a difference in their number does not matter, so that the difference between three and five A's is the same as between three and six. Yet, human perceivers can readily and accurately perceive and discriminate between such arrays. How is this possible? A solution would be to formulate a complementary code that does preserve number. In this chapter an attempt is made to formulate a complementary code of such linear arrays in terms of the spatial ordering of the elements. It results in a unique code for arrays of different length without resorting to a numerical code such as "5 *(A)". A prediction derived from the properties of these complementary codes is that the complexity of arrays of increasing length increases stepwise rather than gradually. This is an interesting prediction as it relates to the observation that perceivers can state the number of elements in small arrays almost immediately, whereas from a certain number of elements upwards they have to count. The above prediction is tested in a same/different task. The subjects are requested to indicate as fast and as accurately as possible whether two successively presented arrays contain an equal number elements.

In Part III, consisting of chapters 4 and 5, the third main issue of this thesis is addressed: the autonomy versus context-sensitivity of the initial organization. In our definition, *primary perception of structure* does not necessarily draw on sources of information other than the stimulus immediately present. Still, this does not exclude the possibility that primary perception can be altered before any stimulation. In chapter 4 this is formulated as the possibility that context may prime the perceptual system during the access phase, rather than contributing only to a post-access phase. In this chapter, many experimental manipulations are combined in an attempt to bias primary perception in favour of one or another organization of a simple line pattern. Subjects are led to expect a particular organization by starting a trial with a special *prime*: a pattern in a segmented (exploded-view) version corresponding to but one of the various possible organizations of the complete pattern. The question is whether this prime, and the expectation it should evoke, changes the initial organization of the complete pattern shown a few seconds later.

In chapter 5 another method is used to investigate autonomy in early perception, in this case with respect to the recognition of spoken words. It has often been stated that speakers of languages such as English and Dutch appear to impose a certain rhythmical structure on their sentences. More technically, their sentences would be "stress timed": The important words, spoken with considerable stress, are assumed to be produced according to a regular temporal organization, or "rhythm" (Martin, 1972). Once locked into the rhythm of a sentence, a listener would be able to anticipate the important words to follow, and to process them more efficiently. We will test the latter role of the global temporal organization for word recognition by artificially speeding up or delaying a stressed word. The subjects have to react as fast as possible upon the occurrence of a specific phoneme placed at the beginning of that word. Known as *phoneme-monitoring*, this task is one of the few that are considered to allow for a measurement of the speed of word-recognition relatively uncontaminated by post-access effects (Cutler & Norris, 1979). Care is taken not to introduce any phonetic discontinuities which would contaminate the results. This is achieved by replacing, in the recorded sentence, the word preceding the stressed word by another recorded word differing in length. If the rhythm creates an effective expectation, this artificial manipulation of the temporal structure should delay the detection of the target phoneme in that word.

Each of the five chapters to follow is written as an independent paper, each with an independent discussion of the results. An on-going overview of the findings, emphasizing the implications for SIT, can be found in the prefaces introducing each Part.

References

Ballard, D. H., Hinton, G. E., & Sejnowski, T. J. (1983). Parallel visual computation. *Nature*, *306*, 21-26.

Buffart, H., Leeuwenberg, E., & Restle, F. (1981). Coding theory of visual pattern completion. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 241-274.

Buffart, H., Leeuwenberg, E., & Restle, F. (1983). Analysis of ambiguity in visual pattern completion. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 980-1000.

Buffart, H. (1987) Seven minus two and structural operations. In E. E. Roskam, & R. Suck (Eds.), *Progress in Mathematical Psychology, I.* Amsterdam: North-Holland Publishing Company.

Bower, G. H., & Glass, A. L. (1976). Structural units and the redintegrative power of picture fragments. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 456-466.

Collard, R., & Buffart, H. (1983). Minimization of structural information: a set-theoretical approach. *Pattern Recognition, 2,* 231-242.
Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In: W. Cooper & E. Walker (Eds.), *Sentence processing.* New York: Lawrence Erlbaum Associates.
Hatfield, G., & Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin, 2,* 155-186.
Helm, P. van der., & Leeuwenberg, E. (1988). Accessibility, a criterion for the choice of visual coding rules. Internal Report 88 NICI 01, University of Nijmegen.
Hochberg, J., & McAllister, E. (1953). A quantitative approach to figural goodness. *Journal of Experimental Psychology, 46,* 361-364.
Koffka, K. (1935). *Principles of Gestalt Psychology,* New York: Harcourt.
Leeuwenberg, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology, 84,* 307-349.
Leeuwenberg, E. (1973). Meaning of perceptual complexity. In D.E. Berlyne & O.O. Madsen (Eds.), *Pleasure, reward, preference.* New York, N.Y.: Academic Press
Leeuwenberg, E., & Buffart, H. (1983). An outline of coding theory: a summary of related experiments. In H.-G. Geissler, H.F.J.M. Buffart, E.L.J. Leeuwenberg & V. Sarris (Eds.), *Modern issues in perception* (pp. 1-47). Amsterdam, NL: North-Holland Publishing Company.
Marr, D. (1982). *Vision.* San Francisco, CA: W.H. Freeman and Company.
Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behaviour. *Psychological Review, 79,* 487-509.
McKay, D. (1950). Quatal aspects of scientific information. *Philosophical Magazin, 41,* 289-301
Mellink, H., & Buffart, H. (1986). Abstract code network as a model of perceptual memory. *Pattern Recognition, 20,* 143-151.
Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision.* New York, NY: McGraw-Hill.
Neisser, U. (1967). *Cognitive Psychology.* New York: Appleton-Century-Crofts.
Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441-474.
Palmer, S. E. (1983). The psychology of perceptual organization. In J. Beck (Ed.), *Human and Machine Vision.* NY: Academic Press.
Reed, S. K. & Johnsen, J. A. (1975). Detection of parts in patterns and images. *Memory & Cognition, 3,* 569-575.
Restle, F. (1982). Coding theory as an integration of gestalt psychology and information processing theory. In J. Beck (Ed.), *Organization and representation in perception.* Hillsdale, N.J.: Lawrence Erlbaum.
Rock, I. (1983). *The logic of perception.* MA: MIT Press.
Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review, 70,* 534-546.

Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proc. Roy. Soc. B., 171*, 297-317.

Tuijl, H. van (1980). Perceptual representation of complex line patterns. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 197-221.

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope & A. Rosenfeld (Eds.), *Human and machine vision.* NY: Academic Press.

PART I: INHERENCE

Preface

Part I consists of only chapter 1. To show that organization is inherent in the initial representation rather than being an epiphenomenon of later stages, a special recognition task is used. The stimulus that has to be recognized, the subpattern, is presented for 10 milliseconds, followed or preceded by another pattern, the complete pattern, with a very short interval in the range of 0 to 60 milliseconds. All contours of the subpattern fall on the same area of the screen as those of the complete pattern. This task is clearly one that yields many sensory effects. These effects have been described in the extensive literature on *masking* (Michaels & Turvey, 1979). Calis (1981, 1983, 1984), however, showed that such a masking paradigm can also reveal early ("microgenetic") phases of perceptual classification (e.g., during face-recognition). In his conception, the first stimulus triggers an hypothesis-testing process. With an appropriate interval between the two stimuli, this perceptual hypothesis is tested on information from the second stimulus. In case of a confirmation, more specific hypotheses can be triggered. In case of a disconfirmation (rejection), the process is disrupted and has to start anew, leading to a loss of information of the first stimulus or, in other words, to the first stimulus suffering from *backward masking* by the second one. Thus, if one has a theory to predict which hypotheses are triggered by the first stimulus and which subsequently are confirmed (or rejected) by the second one, masking effects can reveal states of knowledge "within" perception. This theory will be SIT and the perceptual knowledge that we will look for is knowledge of structural properties of patterns leading to a perceived organization.

As in most studies of this thesis, SIT provides a convenient framework because it is a theory of the representation of perceptual structure. The predictions in chapter 1 are formulated in terms of the *code reference* function that is part of SIT's coding model. Code reference is SIT's way of accounting for the role of context on perception. Code reference enables one to code a sequence of patterns by re-using the structural code of a pattern for the coding of a subsequent pattern. This can result in a lower overall information-load than if both patterns are assigned separate codes. If code reference indeed is more efficient, the two patterns are predicted to receive *compatible* perceptual organizations. For another sequence (in chapter 1 the same patterns presented in the reversed order), code reference can be more expensive than two separate codes; in that case it is not the best (*minimum*) solution to code the second pattern in terms of the first code and the prediction is that the two patterns will lead to *incompatible* organizations. In sum, on the basis of SIT we predict that the interpretation of the second pattern can be different depending on which code of the second pattern yields the lowest overall information-load within the context of the code of the first pattern: Unambiguous patterns can be behave as if their shape is ambiguous.

As will be stressed in the Epilogue after Part III, SIT is a theory of perceptual representations and it does not readily provide a model of the process of perception. SIT does not specify under what circumstances reference to earlier perceived patterns will be made. For one thing, it is obvious that not all patterns ever perceived in an individual's lifetime are equally relevant: There is an upper limit to the effect of context. This issue will be dealt with in Part III. In the first study we will focus on the lower limit: the time needed for the processing of a pattern before it can turn into a relevant context for patterns that are presented later on. As it was formulated in the Introduction, a goal of the first study is to establish the time needed to process the first stimulus from a sensory registration into a representation that implies a perceptual organization. This is also the lower temporal limit to the context effect. If this lower limit turns out to be sufficiently small compared to the latency of reasoning-like processes, it would add to the conclusion that such an experimental method can reveal aspects of the primary perception of structure.

Acta Psychologica 59 (1985) 91-102
North-Holland

KNOWLEDGE WITHIN PERCEPTION: MASKING CAUSED BY INCOMPATIBLE INTERPRETATION *

Emanuel LEEUWENBERG, Lucas MENS and Gé CALIS

University of Nijmegen, The Netherlands

Most people and some scientists assume that knowledge does not affect the perceptual representation of a pattern but merely may change its evaluation afterwards An experiment was set up to show that the perceptual representation itself depends on an interaction with immediate context information, even in the case of clearly unambiguous patterns. In each trial two patterns were presented successively for 10 msec on the same place with an onset asynchrony between 10 and 70 msec Presented in one order, the perceptual representation of the first pattern was predicted to be able to reinforce a latent interpretation of the second one For the same patterns presented in the reversed order, however, no such interaction was predicted Subjects had to identify one of the two patterns of a pair The number of correct identifications differed for the two opposed orders of presentation as predicted The effect was present with stimulus onset asynchronies as short as 30 msec This gives an indication of the rate at which pattern interpretations can be developed

According to Rock (this issue) the perception of a pattern is open to context only if the 'autochthonous' perceptual organization – as processed in isolation – would already incorporate the organization as induced by the contextual bias. In other words, acquired knowledge can only reinforce a latent interpretation. This implies that context can only have an effect if the pattern itself is, at a perceptual level, multi-interpretable or ambiguous. In fig. 1 such a situation is presented. Pattern B in isolation can be interpreted as a square with a line but also as two triangles. Within the context of pattern A however the last interpretation will be preferred.

In the present study an attempt is made to show that even the interpretation of unambiguous patterns may be altered due to the effect of knowledge about context. In fig. 2 both patterns are unambiguous.

* This research was supported by the Netherlands Organization for the Advancement of Pure Research (Z W.O.).

Mailing address: E Leeuwenberg, Dept. of Experimental Psychology, University of Nijmegen, Montessorilaan 3, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

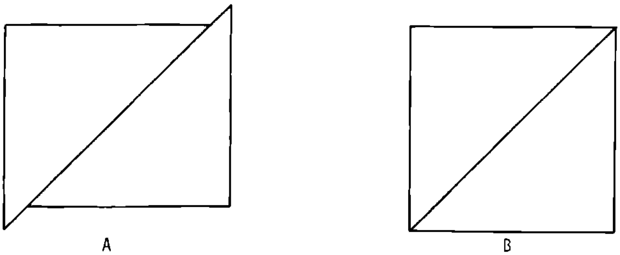


Fig 1 Pattern B by itself is ambiguous it is perceived either as a square plus a line, or as two triangles However, in the context of pattern A, pattern B is organized as two triangles

Pattern A is by preference perceived as a zigzag and pattern W as three triangles (Collard and Leeuwenberg 1981). In the context of A however, subjects may report that pattern W looks like a zigzag plus a line. But of course, simple phenomenological reports are not enough to prove true perceptual effects of context. Response biases or demand characteristics have to be excluded. Therefore we will present stimuli such as pattern A and W tachistoscopically and look for more direct indices of processing. Before full treatment of the experiment, a short digression into the motives and role of this experiment seems useful.

Not only are several scientists proponents of the view that perception is greatly impenetrable by conscious knowledge, but they also hold that perceptual processes are fundamentally different from those underlying logical reasoning or higher cognitive functions in general (Fodor 1983; Kanizsa and Gerbino 1982). We do not disagree with them on the point of impenetrability. However, we do want to show that *within* the perceptual module, states of acquired object knowledge interact in a fashion similar to that during logical reasoning. It is argued here that perceptual representations themselves are constructed under a tendency of the system towards the simplest representation of the given pattern or event (i.e. the minimum principle). It is shown by Buffart (Buffart et al. 1981) that the simplest interpretation has the unique quality of

Fig 2 Pattern W by itself is not ambiguous it is perceived as three triangles Nevertheless, it is shown that in the context of pattern A, pattern W tends to be organized as a zigzag plus a line

accounting for a maximum of all sorts of regularities within the pattern. This is also defended for demonstrations which seem to support Kanizsa and Gerbino's position that patterns are not organized in the simplest way. The present experiment is in line with a series of studies by Calis (Calis and Leeuwenberg 1981; Calis et al. 1983; Calis et al. 1984). In general, he demonstrated that even the most primary stages of perception show reasoning-like characteristics.

As an introduction to the experiment the first of four sets of stimuli which are presented in fig. 3 will be discussed. Set 1 contains six pairs of stimuli (WA, ..., CW). In each pair there is one 'Whole pattern' W from which the other three patterns A, B and C are derived by deleting the same amount of line segments. A, B and C are called subpatterns of pattern W. In the experiment the two patterns of each pair are presented sequentially as depicted in fig. 3. It is important to note that stimulus presentation was not at different places on a screen but such that the two patterns were maximally overlapping. Each pair will now be described in detail.

AW: The sequence AW was already mentioned. The pattern code of A was said to be a good point of departure for the perceptual organization of pattern W. Therefore this sequence is called 'compatible'. A code-theoretical account can be found in Collard and Leeuwenberg (1981). According to their analysis it is simpler to interpret the AW pair as two variants of a zigzag (Information Load = 6) than as the two unrelated entities of a zigzag and three triangles (Information Load = 8).

WA: Pair WA contains the same patterns as pair AW but it evokes a different interpretation. The simplest code of the first pattern W is equivalent to an organization of three triangles. This code, however, is not very useful for coding the subsequent pattern A. It is complicated to transform three triangles into a zigzag. The reason for this is that if a zigzag has to be described departing from a code of the three equilateral triangles, the internal regularity of each triangle in this code has to be abandoned: Each of the base lines which is integrated in the code of an equilateral triangle has to be separated in order to reach a code of a zigzag. Since a common organization is lacking for sequence WA and two distinct interpretations emerge, it is called 'incompatible'.

BW: In the preceding pair WA, the incompatible order started with the

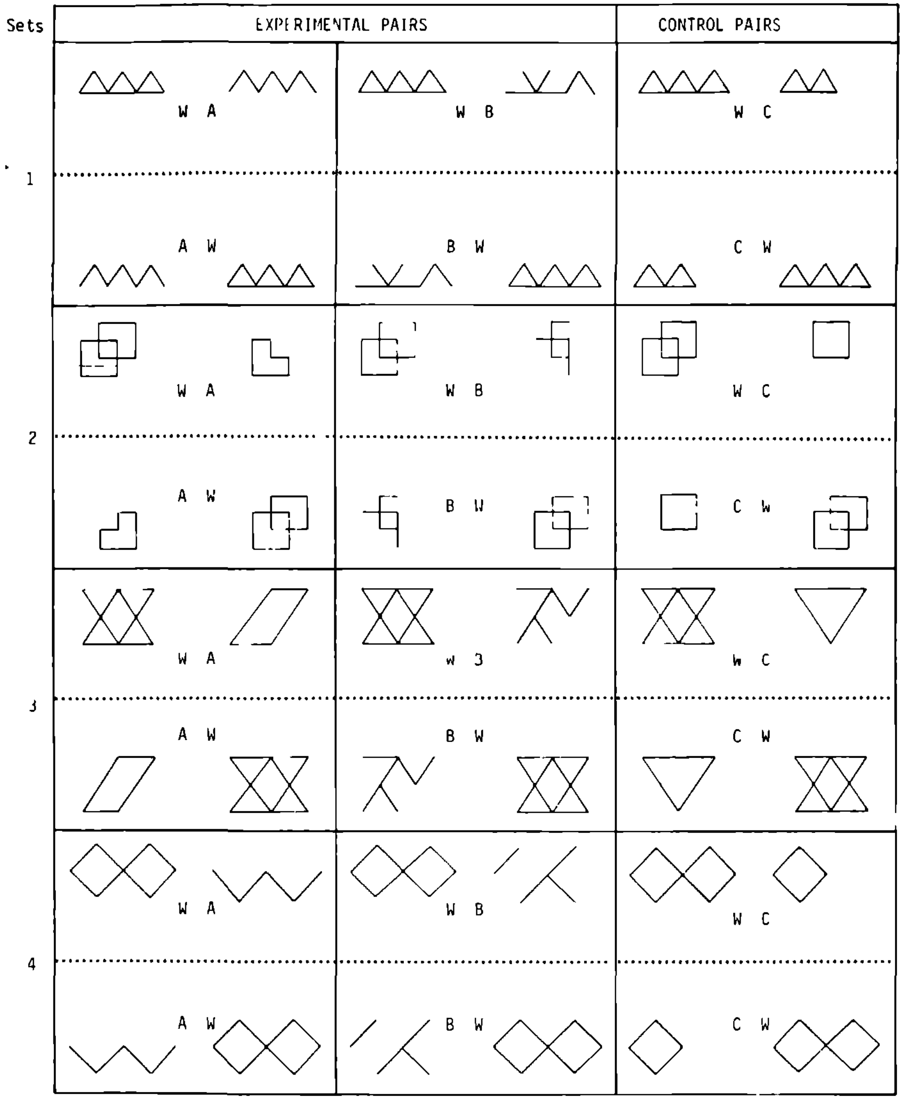


Fig 3 Pairs of patterns which were subsequently presented in the experiment In each of the 4 sets there are three pairs in two opposed orders In the first two experimental pairs the relation between the pattern organizations shows an asymmetry WA is incompatible, AW is compatible, WB is compatible and BW is incompatible Both WC and CW are compatible

whole pattern. In BW, also incompatible, the whole pattern comes last. Here pattern B with its peculiar V-shapes is of little use for the three triangles which follow.

WB: Starting with three triangles, the subsequent complex pattern B can now be profitably perceived as three triangles lacking some line segments. Therefore this sequence is called 'compatible'.

CW and WC: In contrast with the preceding pairs this combination of patterns is compatible in both orders of presentation. Both are interpreted by preference as triangles irrespective of order. Below it will be shown that CW and WC are control sequences.

In order to be able to list our predictions it is necessary to give a brief indication of the procedure. Subjects were shown a pair of patterns, each of which was presented for 10 msec with varying onset intervals between both patterns in the range of 10 to 70 msec. After that, subjects had to identify the subpattern. After each pair of flashes all subpatterns included in the experiment were shown as response alternatives.

The general prediction is that the number of correctly recognized subpatterns depends upon the compatibility between pattern interpretations. As compatibility can be varied using the same pictures (eg. AW versus WA), a number of other effects on recognition is controlled for. One important factor remains: Later patterns of a pair are prone to be more easily identified than the first ones, due to a recency effect (Michaels and Turvey 1979). Therefore, control pairs WC and CW which do not differ in compatibility are introduced. As a result of recency we predict that if C is the later member of a presented pair it will be recognized more often than when it is the first one shown. This effect of recency leads to the following prediction: $wc > cw$. In this equation lower-case letters represent the number of correct identifications in the corresponding conditions.

The general prediction above can be given more vividness by a tentative model of the effect of compatibility in the four experimental pairs. In a previous study using the same presentation of stimuli (Mens 1983), it was shown that compatible pairs of the type WB were relatively often perceived as two separate patterns and the incompatible pairs BW as one single pattern. We suppose that first patterns in such incompatible pairs are masked by the second pattern. This is in line with the commonly found backward (type B) masking (Michaels and

Turvey 1979; Bachmann and Allik 1976) from which first stimuli suffer. In fact, the clerk-customer metaphor of backward masking (Turvey 1978), with its explicit role of a central processor's attention strategy, seems a good perspective from which to appreciate the present account of masking. In these studies, stimuli were used which are generally incompatible by our definition. We predict that subpatterns in BW pairs will be difficult to recognize. In our notation: $wb - bw > 0$. For the incompatible pair WA we expect a suppression of the first pattern WA. We assume that this will enlarge the 'Pragnanz' of the second pattern A. This assumption as such is not supported by any accepted evidence, but we take this facilitation as a natural consequence of the replacement of one interpretation by a very distinct other one due to incompatibility. Thus we predict: $wa - aw > 0$. The general prediction can be formulated as:

$wa - aw \neq wc - cw$, and

$wb - bw \neq wc - cw$.

Our account of specific effects would yield:

$wa - aw > wc - cw$, and

$wb - bw > wc - cw$.

Experiment

Method

Material

Sixteen line drawings as presented in fig 3 were copied on slides, resulting in bright figures against a dark background Two Random Access projectors were positioned so that the two patterns of a pair overlapped maximally Projection was slightly out of focus in order to compensate for small imperfections in overlap The visual angle subtended by the patterns was about 7 degrees The experimental room was dimly lit In fig 4 the lay-out of the response panel is given. All subpatterns are depicted above the corresponding response keys A starting key is provided at the lower left side Next to each set of three subpatterns is a 'don't know' key which could be chosen in case of severe doubt

Procedure

Thirty students of the University of Nijmegen, most of them in their first year, served as *S*s Each *S* was tested individually After a training session of thirty trials

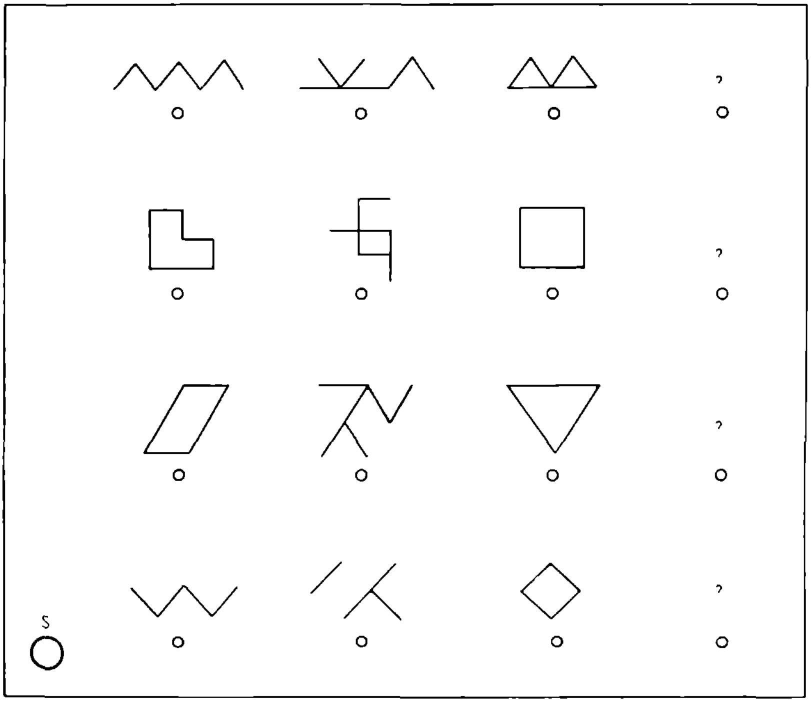


Fig. 4. Response panel. The task was to identify the subpattern in each presented sequence and to indicate the answer by pushing the corresponding key The keys in the right column are used in case of severe uncertainty. 'S' is a self-packed starting key

two experimental sessions followed of about thirty minutes with a short break in between. Trials were self-initiated. After the starting key was pressed, immediate presentation of the two patterns (10 msec each) followed with one of 7 different onset asynchronies (SOA). These SOA's varied between 10 and 70 msec with steps of 10 msec. A trial was completed when the *S* selected one response key. The instruction was to detect which of the 12 possible subpatterns had been 'flashed' together with the whole pattern. There was a time limit of 10 sec but no emphasis on speed. *S*s were run twice through the complete series. This resulted in $3 \times 4 \times 7 \times 2 = 168$ trials arranged in one random order.

Results

In fig. 5 the percentages of correctly identified subpatterns are presented as the average of the four sets shown in fig. 3. The base axis represents SOA. Above and below the curves the types of stimuli are depicted which gave rise to these curves. All upper curves refer to sequences with the subpattern last, all lower ones with subpatterns first. This indicates the general effect of recency.

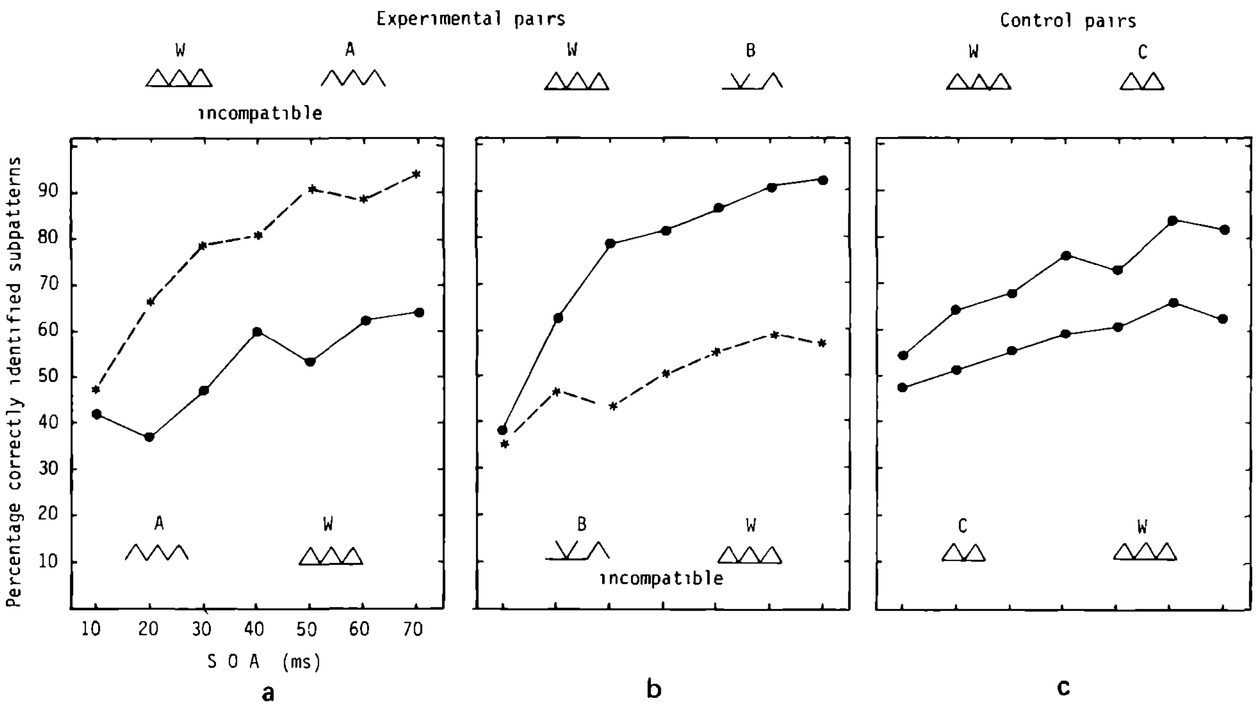


Fig 5 Results averaged over the 4 sets of fig 3 only the patterns of set 1 are depicted The curves reflect the percentage correct identifications of the subpatterns for several stimulus onset asynchronies The difference between the upper and the lower curve for each experimental pair (a b) is larger ($p < 0.001$) than for the control pair (c) Hence an effect of compatibility of pattern interpretations is obtained

Curves marked with asterisks indicate incompatible sequences. The larger gaps between curves in figs. 5a and 5b relative to the gap in fig. 5c can be interpreted as the effect of compatibility. In other words, the results are in agreement with the general prediction.

Analysis of variance showed the following:

$$wa - aw < wc - cw, \quad F(1, 29) = 22.88, \; p < 0.0001 \quad \text{(effect in fig. 5a)};$$

$$wb - bw > wc - cw, \quad F(1, 29) = 17.96, \; p < 0.0002 \quad \text{(effect in fig. 5b)}.$$

The two analyses on the average results described above were performed for each one of the four sets of fig. 3. The general trend was preserved in all eight cases but in three cases no significance was obtained ($p > 0.05$). These three are: set 2 – column 1 and 2; set 3 – column 2.

Analyses were performed on the overall results for each SOA separately. The effect of compatibility in fig. 5a was significant at SOA's of 20, 30, 50 and 70 msec ($p < 0.01, 0.001, 0.001, 0.05$). In fig. 5b all SOA's larger than or equal to 30 msec yielded significant surplus ($p < 0.05, 0.001, 0.01, 0.01$ and 0.01). It can be concluded that the predicted effect of incompatible contexts occurs with SOA's as short at 30 msec.

Alternative approaches

There are five stimulus-features which could have been alternative sources of the effects found above. They are all more or less confounded with the property of compatibility.

Firstly, the control patterns WC and CW of fig. 3 are all closed figures. Indeed, pairs consisting of closed complete patterns as well as closed subpatterns showed less

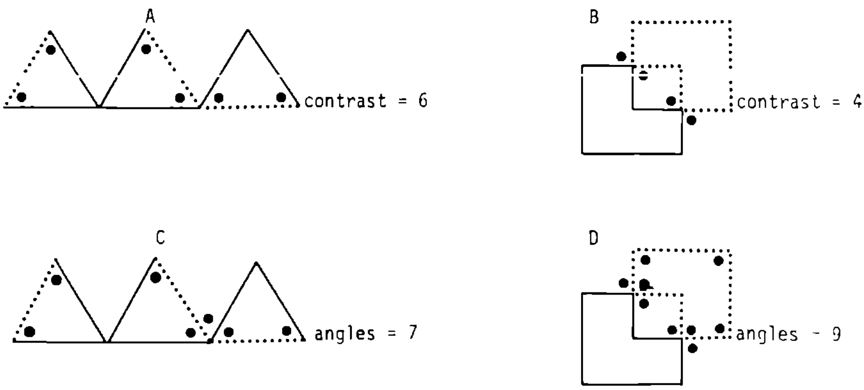


Fig 6 Four examples of calculation of contrast- and angle-gradients. The dots in pattern A and B indicate the points of contrast between subsequent patterns in an experimental sequence The dots in pattern C and D indicate the amount of angles which appeared or disappeared.

effect of order than pairs with open subpatterns ($F(1, 29) = 18.50$, $p < 0.0002$). However, no such interaction of order of presentation with closure is found *within* the first column of fig. 3 (without the confounding variable of compatibility). Furthermore, closure of subpatterns did not have a main effect on their identification ($F(1, 29) = 2.56$, $p < 0.1206$).

Secondly, the existence of a feature-detector which would be sensitive to contrast gradients at junctions can be assumed. In figs. 6A and 6B the dots indicate the points of change in contrast. However, the number of contrast points was not related to total percentage correct identification, nor did it interact with order of presentation.

Furthermore, no effect was found of the amount of angles which (dis)appeared between the two stimuli (see figs. 6C and D). No effect was found of the complexity of the separate patterns as specified by a structural information measure (cf. Buffart et al. 1981), nor of an estimated spatial frequency gradient.

Discussion

The main conclusion is that the perceptual organization and subsequent identification of a pattern can be drastically altered within the context of another one immediately preceding it, even when both patterns clearly are not ambiguous as such.

This context effect is not anchored in a special property of the stimuli but is dependent upon the perceptual tendency towards the simplest interpretation: one which accounts best for the regularities within the patterns. Even unambiguous patterns are interpreted according to what at any moment is computed to be the most economical representation of the given stimuli. Thus even unambiguous patterns are multi-interpretable. Therefore, within perception, special kinds of knowledge are at work: knowledge about structural properties of patterns and knowledge in the form of rules to manipulate and combine these properties.

The appearance of the effect of compatibility with SOA's as short as 30 msec deserves separate discussion. This supports that such a short time of extra processing of the first stimulus is sufficient to produce a code which represents the first pattern as distinct from what comes after it. In other words: the temporal resolution of the central pattern processor seems to be as small as 30 msec with simple stimuli as these. However, this conclusion is one out of several possible ones. For instance, there are about 100 msec involved in the transfer of retinal information to the central brain. During this time lag, any sort of interference may have occurred, which might delay further processing.

We assume, however, that these time lags are equal for the different stimuli. It is also possible that the two subsequent patterns in each trial are stored in a buffer in the given order. If so, the pattern interpretations might be built up in a time far beyond 30 msec central processing, and the present data would be explained by nonperceptual selective forgetting instead of visual masking. In our opinion, this presentation of the matter requires a train of assumptions far less parsimonious than the one we proposed in order to explain the results. Given such a complete storage, why would the incompatible sequence not be coded in the reverse – compatible – order? And, putting it more strongly, how can the relation between masking and SOA be explained?

If one accepts that a first pattern code is active from about 30 msec onwards, a final point can be made about the position of the minimum principle. These 30 msec hold equally for the W patterns in this experiment as for the B patterns. B patterns are more complex than W patterns. Thus it seems that there is no compelling relation between the complexity of a pattern and the duration of encoding. On the other hand, embeddedness experiments show two things (van Tuyl and Leeuwenberg 1980). Firstly, the detection of a subpattern in a larger pattern requires a long time if the description of the embedding pattern according to its minimal organization is simpler than the description in terms of the subpattern. Secondly, these search times were in general much longer than 30 msec. Thus one is forced to conclude that the excess time needed to perform a difficult embeddedness judgment cannot be simply accounted for by assuming a longer primary processing stage. In our opinion, reaction times in tasks such as these reflect decision processes and a test phase of the preferred code. This is supported by double-stimulation experiments done by Calis and Leeuwenberg (1981) and in agreement with experimental findings of Boer (1982). He concludes the global pattern aspects are dominant with respect to details during a code test phase but not during the encoding process itself.

References

Bachmann, T. and J. Allik, 1976. Integration and interruption in the masking of form by form. Perception 5, 79–97.

Boer, L.C., 1982. Perception of a visual stimulus in a selective attention paradigm. P.E.D. report, Vrije Universiteit Amsterdam.

Breitmeyer, B.G., M. Rudd and K. Dunn, 1981. Metacontrast investigations of sustained-transient channel inhibitory interactions. Journal of Experimental Psychology: Human Perception and Performance 7(4), 770–779.
Buffart, H., E. Leeuwenberg and F. Restle, 1981. Coding theory of visual pattern completion. Journal of Experimental Psychology: Human Perception and Performance 7, 241–274.
Buffart, H., E. Leeuwenberg and F. Restle, 1983. Analysis in visual pattern completion. Journal of Experimental Psychology: Human Perception and Performance 9(6), 980–1000.
Calis, G.J.J. and E.L.J. Leeuwenberg, 1981. Grounding the figure. Journal of Experimental Psychology: Human Perception and Performance 7, 1386–1397.
Calis, G.J.J., J.M. Sterenborg and F.J. Maarse, 1984. Initial microgenetic steps in single glance face recognition. Acta Psychologica 55, 215–230.
Calis, G.J.J., H.L. Teulings and P.J.G. Keuss, 1983. In search of writing and reading habits in the microgenetic phase of letter recognition. Acta Psychologica 54, 313–326.
Collard, R. and E. Leeuwenberg, 1981. Judged temporal order of visual patterns. Canadian Journal of Psychology 35, 323–329.
Fodor, J.A., 1983. The modularity of mind. Cambridge, MA: MIT Press.
Kanizsa, G. and W. Gerbino, 1982. 'Amodal completion: seeing or thinking'. In: J. Beck (ed.), Organization and representation in perception. Hillsdale, NJ: Lawrence Erlbaum.
Leeuwenberg, E., 1971. A perceptual coding language for visual and auditory patterns. American Journal of Psychology 84, 307–349.
Leeuwenberg, E., 1982. The perception of assimilation and brightness contrast. Perception and Psychophysics 32(4), 345–352.
Leeuwenberg, E. and H. Buffart, 1984. The perception of foreground and background as derived from structural information theory. Acta Psychologica 55, 249–272.
Mens, L., 1983. Pattern code in construction: the microgenesis of perceptual interpretation. University report. K.U.N. University of Nijmegen.
Michaels, C.F. and M.T. Turvey, 1979. Central sources of visual masking: indexing structures supporting seeing at a single, brief glance. Psychological Research 41, 1–61.
Restle, F., 1982. 'Coding theory as an integration of Gestalt psychology and information processing theory'. In: J. Beck (ed.), Organization and representation in perception. Hillsdale, NJ: Lawrence Erlbaum.
Rock, I., 1985. Perception and knowledge. Acta Psychologica 59, 3–22.
Turvey, M.T., 1978. 'Visual processing and short-term memory'. In: W.K. Estes (ed.), Handbook of learning and cognitive processes, Vol. 5. Hillsdale, NJ: Lawrence Erlbaum.
Van Tuyl, H.F. and E. Leeuwenberg, 1980. Perceptual interpretation of complex line patterns. Journal of Experimental Psychology: Human Perception and Performance 6, 197–221.

PART II: COMPLETENESS

Preface

From Part I we conclude that the experimental method presently used is adequate to show early processing effects that can be attributed to the primary perception of structure. Because of this method, any explanations in terms of sensory masking, local features, or as in terms of biases due to cognitive reconstructions of the stimuli are implausible.

On the basis of SIT codes, we were able to predict in which cases the representation of the first stimulus is an effective context for the interpretation of the subsequent pattern, and in which cases not. From the results we concluded that even the organization of unambiguous patterns may to some extent depend on context. However, there is one aspect that needs to be clarified. The predictions of the preceding study were based on a comparison of the information loads of SIT codes of the sequence with and without *code-reference*. Although code-reference is only a technical means to derive predictions (SIT not being a process theory), it suggests a perceptual process different from the sequential hypothesis-testing model adopted by Calis (see the Preface to Part I). Does this comparison correspond to an actual perceptual choice between two alternative representations of the stimulus sequence? This can not be concluded from the results of the preceding study. The (in)compatibility of the two patterns can perceptually be established in a purely sequential fashion (cf. Calis & Leeuwenberg, 1981): A hypothesis triggered by the first stimulus tested on sensory information of the second stimulus. It is possible that confirmation of this hypothesis does not depend on the adequacy of concurrently available alternative hypotheses. Instead, confirmation can result from some unknown satisfaction-criterion being reached. In the study reported in chapter 2, an attempt is made to show that primary perception does entertain concurrently available alternative pattern-organizations. In other words, the first study showed that unambiguous patterns are multi-interpretable; the following study shows that they are also ambiguous in the sense that alternative organizations are present in perception, "hidden" beneath the dominant organization. To this end, the method of chapter 1 is refined in chapter 2 by choosing specially selected pairs of a complete pattern and a subpattern. The subpatterns are closely matched with respect to their structural complexity. More importantly, one of the subpatterns corresponds to the preferred organization of the complete pattern and a second subpattern to one that is almost as good: a second-best organization of which the SIT code has the property of complementarity. By also applying a response-bias correction, we hope to be able to show the hidden concurrent presence of the second-best interpretation.

On a more theoretical level, the goal of the chapters in Part II is to find support for an important implication of adopting a coding model such as SIT: As has been argued in the Introduction, for most patterns there is not a single code that covers the structural regularities completely. Therefore, SIT predicts that for an adequate or *complete* representation, more than one perceptual organization is required. So again experimental evidence of the concurrent presence of a hidden alternative organization is required. In addition, SIT's measure of the perceptual strength of alternative organizations is based on the ratio of the corresponding codes, suggesting concurrent presence once more.

In the Introduction structural and metrical information have been opposed as two different aspects of patterns. Structural codes were described as being possible only if one abstracts from metrical differences. The central question of chapter 3 is whether the completeness of representation is only relevant for purely structural properties of a pattern. Therefore a somewhat more metrical type of representation will be investigated. The task of the subjects is to maintain for a short period an exact impression of the number of identical elements in a row in order to compare it with the number of elements in a second row. Such a *numerosity judgement* obviously is not representative for purely metrical tasks (such as weight estimation), but it does require a completeness of representation that is different from the structural one of chapters 1 and 2. Therefore, a specific coding model is proposed. Instead of replacing identical elements by one symbol (as in a SIT code), the relative position of pairs of elements is completely specified, This insures that each element is represented distinctly from the others, which implies that the aspect of number is represented completely.

Journal of Experimental Psychology:
Human Perception and Performace
1988, Vol. 14, 4.

0096-1523/88/$.00.75

Hidden Figures Are Ever Present

Lucas H. M. Mens and Emanuel L. J. Leeuwenberg
University of Nijmegen, the Netherlands

Preference judgments about alternative interpretations of unambiguous patterns can be explained in terms of a rivalry between a preferred and a second-best interpretation (cf. Leeuwenberg & Buffart, 1983). We tested whether this second-best interpretation corresponds to a suppressed but concurrently present interpretation or whether it merely reflects an alternative view that happens to be preferred less often. Two patterns were presented immediately following each other with a very short onset asynchrony: a complete pattern and one out of three possible subpatterns of it, corresponding to the best, the second best, or an odd interpretation of the complete pattern. Subjects indicated which subpattern was presented by choosing among the three subpatterns shown after each trial. The scores, corrected for response-bias effects, indicated a relative facilitation of the second-best interpretation, in agreement with its predicted "hidden" presence. This result is more in line with theories that capitalize on the quality of the finally selected representation than with processing models aimed at reaching one single solution as fast and as economically as possible.

GENERAL INTRODUCTION

In general, we are not confused about the shape of things. Our knowledge of natural objects as well as the overwhelming redundancy in the visual input helps us to establish one firm shape interpretation at a time. Artificial pat-

---

This research was supported in part by the Netherlands Organization for the Advancement of Pure Research (ZWO).

The authors wish to thank Noud van Kruysbergen for his assistance in these experiments and J. G. W. Raaijmakers for the realization of the bias model. Several readers of an earlier draft of this paper have provided thoughtful remarks. Special thanks are due to William P. Banks and Julian Hochberg for a number of truly essential suggestions. The title of this paper was discovered by Mary Jones.

Correspondence concerning this article should be addressed to Lucas H. M. Mens, P.O. Box 9104, 6500 HE Nijmegen, the Netherlands.

terns can be given alternative interpretations but often only under prolonged viewing conditions (for example, when searching for the hidden face in a drawing of a tree). Some very unusual patterns, specially designated as ambiguous ones, are seen in two ways spontaneously, but not *simultaneously*. Nevertheless, below we will refer to some aspects of the perception of *unambiguous* patterns that may be explained by assuming the simultaneous presence of a second, "hidden" interpretation. "Ambiguous" patterns, in this view, are special only because the alternative interpretations are of about equal strength and therefore readily perceived at least some of the time. Our aim is to find experimental evidence which may help to decide between two options: (a) Alternative shape interpretations of any pattern are generated one after the other, when they are generated at all; and (b) more than one shape interpretation is concurrently present, even when an unambiguous pattern is perceived.

These options may be put in a broader perspective that includes two opposing views on perception. The first view emphasizes aspects of the process of perception and holds that, for instance, its speed and economy favor Option 1. The opposing view focuses upon the structural quality of a final representation. First some phenomena and views favoring the first option will be given.

Option 1. Processing Minimization and Subsequent Interpretations

A relevant issue is the phenomenon of reversible figures. The two ways to see the Necker cube, for instance, seem to be incompatible. Only one is experienced at each moment. A straightforward explanation of this phenomenon relies on the assumption that the underlying representations are also incompatible and become active only one at a time. This assumption underlies one line of models built around the concept of neural satiation (Orbach, Ehrlich, & Heath, 1963; Kawamoto & Anderson, 1985). Satiation causes a decay of the sensory support for the interpretation "in residence," followed by a shift into another global state of the system representing the alternative interpretation. Kawamoto and Anderson used a network model as a natural means to formalize such a parallel process. The appeal of network models in current perceptual theories is related to the general idea that perception minimizes processing costs by instantiating a locally cooperative hill-climbing procedure (Attneave, 1982) or a similar massively parallel system (for a review, see Ballard, Hinton, & Sejnowski, 1983). This class of models seems to be compatible with properties which are frequently ascribed to perception, namely, that (a) it is not confused by general cognitive convictions; (b) it is computationally shallow, avoiding complex inferences; and (c) it is based on reliable visual features. These properties explain why perception is extremely fast. Furthermore, the computability of the envisaged procedures increases the plausibility of such approaches.

As a typical consequence of processing minimization, one would not expect the concurrent presence of other representations of a pattern besides the dominant one. For example, Hatfield and Epstein (1985, p. 162) state that "Attneave's formulation suggested that the minimum principle guides the construction of a single (most economical) representation ...." In a similar vein, Perkins (1983) remarks that: "minimal search suggests that the human perceiver proceeds along paths of highly reliable recognitions and inferences in order to avoid backup or construction of parallel alternatives." (p. 357) and: "the human procedure seems to pursue only one [interpretation, LM] at a time, attempting another only if the first fails, as in case of viewing the Penrose triangle" (p. 358). An extreme position with respect to this issue is represented by the Gibsonian direct pick-up of invariants as it does not imply the concept of interpretation in the first place, let alone that of concurrent ones.

Option 2. Structural Optimization and Concurrent Interpretations

As early as 1850, Herbart proposed that the degree to which one interpretation of an object dominates other ones is the result of an overall competition between all representations of that object generated by the visual system. He proposed an "easy" differential calculus which formalized such a mechanism. Although Herbart intended this idea primarily for simple sensations, his theory gives a clear demonstration of the idea that (a) the conscious percept is but one out of a host of subliminal candidates, and (b) that the dominant percept may vary in strength depending on competition raised by alternative representations. The first aspect is given a somewhat different formulation in theories of perception which assume that the preferred interpretation involves a *selection* out of alternative ones. The likelihood principle (von Helmholtz, 1867/1962) is an account of perception with a "selective", hypothesis-testing quality to it. Gregory and Gombrich (1973), a modern proponent of the likelihood principle, states about the Necker cube reversal:

> Spontaneous change implies that not only must there be in operation the current or reigning hypothesis—which is the present perception—but also there are more or less ready-formed rival hypotheses waiting to challenge and overthrow the reigning hypothesis (p. 93).

Another selection criterion is the *minimum principle* (Hochberg & McAlister, 1953; see also Hatfield & Epstein, 1985). In studies based on the minimum principle in the context of a coding model (Leeuwenberg, 1971), both the presence of the hidden interpretation and the competition it raises to the preferred one play central roles. Again, perceived ambiguity is seen as a manifestation of a more general principle of subliminal competition. The same rivalry also has been used to explain phenomenal differences between variants of unambiguous patterns. A demonstration is given in Figure 1, which

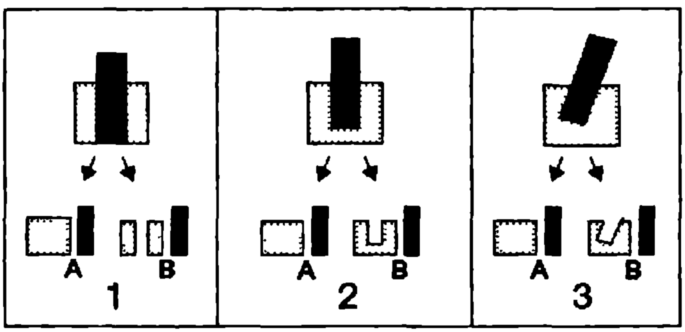


Figure 1. The strength of the preferred occlusion interpretation (A) increases from Interpretation 1 to 3; this can be related to the decreasing simplicity of the alternative mosaic interpretation (B).

shows two possible interpretations (A and B) for three variants of a pattern. The occlusion interpretation (A) is clearly preferred for Variant 3 in Figure 1, is somewhat less clear for Variant 2 and is considerably less for Variant 1.

Note that Organization B is simplest in 1, less simple in 2 and most complex in 3, while Organization A is the same in all three cases. Preference judgments for this phenomena and for a variety of others (for an overview see Leeuwenberg & Buffart, 1983) have been explained on the basis of the ratio between the simplicity of two formal descriptions (or "codes"): the simplest one (A), which captures a maximum of figural regularities, and the second best (B), which expresses regularities not covered by the simplest representation (in the case of Figure 1, the mosaic interpretation). The second-best code may, therefore, be called "complementary" to the best code. However, one might also argue that such preference judgments involve prolonged processing and subtle comparisons of subsequently generated interpretations (Option 1). Our aim in this report is to find more direct evidence for the *concurrent* presence of an alternative interpretation (Option 2). Given the importance of the rivalry between the preferred and second-best interpretation referred to above, we think it justified to search for the hidden presence of that second-best interpretation in particular. In effect, we will contrast the presence of the second-best interpretation with that of a third, rather unlikely, interpretation which may be assumed not to be generated at all (but see the General Discussion).

Concurrent Activation of Word Meanings and Priming

The idea of simultaneously activated interpretations has gained more attention in the linguistic area of word recognition than in visual perception. Reviewing theories about lexical ambiguity, Simpson (1984) stressed that "word-level ambiguity may be viewed as a critical, special case of more general word recognition processes" (p. 316). The interpretation of a word, even one that is judged to be unambiguous, is concluded to be the result of an initial activation of many (perhaps all) possible meanings constrained by frequency and/or sentence context. One of the most frequently used methods to establish initial word-ambiguity is that of the semantic priming task. Characteristic for such a task is that a lexical response (for example naming) must be given to a test word which is preceded by an ambiguous word (which, in turn, is preceded by a disambiguating context). Response facilitation is taken to indicate that the test word matches one of the meanings of the ambiguous word. Thus, with the test word one "probes" which interpretations are activated by the preceding stimulus (Conrad, 1974; Neely, 1977). By varying the time between the ambiguous word and the test word, one typically finds that the initial ambiguity disappears within a fraction of a second (Tanenhaus, Leiman, & Seidenberg, 1979). It may be clear that a similar method for visual patterns would be appropriate in order to decide between Options 1 and 2. Below, we will introduce such a method.

Probing Visual Interpretations

In several studies the relative goodness of alternative interpretations has been investigated using a *part-probe* or *embedded-figures* detection task: Subjects have to decide whether or not a particular pattern is part of a "complete" pattern presented before or after it (Gottschaldt, 1929; Reed & Johnsen, 1975; Palmer, 1977; Van Tuyl, 1980). As a rule, a hierarchy of good, reasonable and bad parts can be obtained instead of a strict dichotomy between, for instance, one good part corresponding to the preferred organization and all other parts which are equally bad. However, the methods used (phenomenal report, reaction times) do not exclude the effect of postperceptual processes of a problem-solving kind. More strictly controlled are *object superiority* studies which show facilitated detection of a tachistoscopically presented small part when it is followed by a second line drawing, which together add up to a structured (3-D) object (McClelland & Miller, 1979; Weisstein & Maguire, 1978). In the experiments to be reported, we use a part-probe task with very short stimulus onset asynchronies (SOAs) in order to minimize the effects of extended postperceptual classifications. Furthermore, a forced-choice recognition response is used which enables the application of a strict control on response bias effects.

Figure 2 shows one of the complete patterns of the experiments. Several segmentations can be made for this pattern. Segmentation A corresponds to the preferred interpretation for most viewers, and B corresponds to the second best, while C represents an unlikely interpretation of the complete pattern.

We probed for the presence of interpretations by examining recognition of a subpattern (S) which will be presented after the complete, or whole, pattern (W) with a short SOA (Condition W-S). The position of the subpattern matched the position of the corresponding part within the complete pattern. In other words, when presented simultaneously, the subpattern would fuse completely with the complete pattern. Recognition scores were established by presenting subpatterns A, B, and C after the fast presentation of the com-

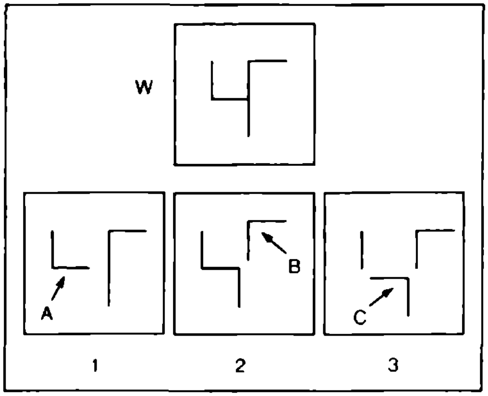


Figure 2. A complete pattern (W) and three subdivisions representing the best (1), the second best (2) and an unlikely interpretation (3). (The visibility of subpattern A, B, or C presented shortly after the complete pattern [condition W-S] is assumed to reflect the concurrent presence of the corresponding interpretation. Note: Complete pattern [W] and subpattern [S] are presented with maximal overlap.)

plete pattern and the subpattern and by asking the subject to indicate which of these three was presented in the sequence. We discerned four factors contributing to the recognition scores.

1. Inhibition. Because of the short SOA, the recognition of the subpattern was expected to be more difficult than if presented without the complete pattern. Therefore, we use the term *facilitation* only in a relative sense—that is—only to refer to a relative advantage or disadvantage for one of the subpatterns over the others.

2. Response bias. Complete patterns are, by definition, larger and thus more conspicuous than the subpatterns. We expect that the complete pattern would induce a considerable response bias irrespective of the subpattern that was actually displayed. In fact, we established in three ways (see Experiment 1) that the organizations related to subpatterns A, B, and C are preferred in the order: A > B > C. Complete patterns that did not fulfill this requirement were excluded from the analysis. Thus, we expected a similar response bias. To accommodate this, we developed a special correction method based on

the distribution of incorrect responses to uncover the "true" recognition of the subpatterns.

3. Preference. Apart from inducing a postperceptual response bias, differences in the preference for the three organisations may also be immediately effective and change the true, bias-corrected recognition of the corresponding subpatterns. If so, we expected it to result in a decreasing score: A > B > C.

4. Concurrent presence. The main goal was to discern the effect of a hidden interpretation if it was concurrently present. Drawing from semantic priming studies, we expected the recognition of a subpattern to be facilitated if it fits into one of the interpretations of the complete pattern which is active, be it the preferred one or the second best. Two possible assumptions concerning the presence of the second-best interpretation can be distinguished, corresponding to the two options introduced above: it is either absent (Option 1) or present (Option 2). These possibilities may be symbolized as: A(1) B(0) C(0) and A(1) B(1) C(0) respectively (note that C[0] in both cases holds under the assumption that Interpretation C is not accessible). Thus, any facilitation of the recognition of subpattern B relative to that of subpattern C may be seen as an indication of the concurrent presence of the second-best interpretation, in line with Option 2.

Unfortunately, an intermediate score for subpatttern B has two reasonable alternative explanations. First, we already mentioned that preference may give similar results. Second, the recognition scores will necessarily be averages over subjects and trials. It has to be taken into account that the Subpattern B may yield an intermediate score, not because it is concurrently present and hidden, but because it is the only one present on a minority of trials (and perhaps even for some subjects only). This is the third possibility concerning the presence of the subpatterns, roughly symbolized by: A(1) B(1/2) C(0). We refer to it as a *probabilistic presence*. It is inspired by the observation that the proportion of time an alternative configuration of a multistable pattern is perceived is a direct function of the probability that it will be perceived initially (Sadler & Mefferd, 1970). We will evaluate these two alternative explanations by comparing the scores for the W-S order of presentation (complete pattern first and subpattern second, the main condition) with the scores for the reversed order of S-W. The relative preference for the corresponding interpretations of the complete pattern can be assumed to be equal in the two conditions. As for probabilistic presence, the idea is that in the S-W condition, the subpatterns are initially interpreted as being distinct objects, without any specific relation to the complete pattern and therefore without an effect on its interpretation. A probabilistic presence would then result in the same relative frequencies of occurrence of the alternative interpretations as in the W-S condition: A(1) B(1/2) C(0). Therefore, it would also not predict any specific difference between S-W and W-S.

In sum, of the four factors, only one is predicted to result in an extra facilitation of subpattern B when it is presented after the complete pattern, relative to the reversed (S-W) condition: the *concurrent presence* of the second-best interpretation (Option 2). This will become evident as an interaction between subpatterns and order of presentation.

EXPERIMENT 1

Method

Material. In the following, we will use the term *W-pattern* to refer to a complete pattern and the term *S-pattern* instead of *subpattern.* Four sets of W-patterns and S-patterns are depicted in Figure 3. From each W-pattern three S-patterns are derived.

The S-patterns contain an equal number of line lengths (e.g., five elements in the first set and two elements in the fourth set of Figure 3) and are meant to be of about equal complexity. In a control condition *Sub-only*, complexity indeed appeared to be balanced: The recognition of an S-pattern A, B, or C, followed by an overlapping mask (a rectangular grid), differed only marginally (see Table 1).

The S-patterns are intended to be compatible with the preferred interpretation (A), the second-best interpretation (B), and a very unlikely interpretation (C). This relative order is established in a pilot study where 10 subjects were given three copies of each W-pattern and were asked to produce three phenomenal subdivisions of the W-pattern (best, second best, third best) by marking some of the segments. The segmentation corresponding to the S-pattern of Type A was produced by nearly all subjects as the best segmentation for each of the W-patterns; S-patterns of Type C were indicated in only a few cases. S-patterns of Type B yielded intermediate scores for a majority of the subjects. The four sets of Figure 3 are selected out of eight sets, because for these four the three S-patterns were chosen with the same rela-

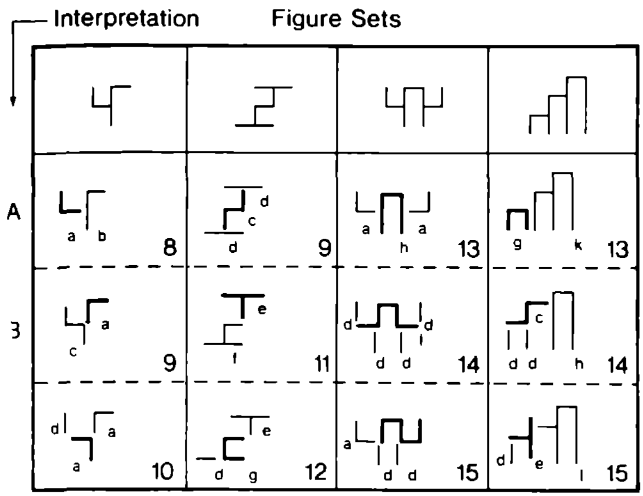


Figure 3. The four figure sets of a complete pattern (top) and three subpatterns used in the experiments. (Subpatterns are indicated by fatter lines in this figure only. The subdivisions in row A, B, and C correspond to the preferred, the second best and an unlikely interpretation of the complete pattern. A coding model has been applied to obtain complexity measures of the subdivisions [bottom right hand corner]. Complexity is calculated as the number of descriptive parameters of the structural codes of the components [letters, see Appendix]).

tive preference (A > B > C) in a control condition consisting of a tachistoscopic presentation of the W-pattern without an S-pattern. Furthermore, the relative preference is also consistent with the complexity of subdivisions of the W-patterns as calculated using the coding rules proposed by Leeuwenberg and Buffart (1983, see Appendix). This amounts to three independent tests which confirmed the order of preference between the three S-patterns for the four W-patterns that were selected. Four other sets failed to produce a consistent relative preference and were therefore discarded from the analysis.

Wirelike patterns are chosen because surfaces—although they are probably more common stimuli—would introduce a number of confounded aspects. It seems very difficult to construct three surface segmentations of the same complexity and the same area.

The patterns were computer generated on a display (Vector-General Display Series 3, Model 2D3 with P4 phosphor), positioned at a distance of 2 m in front of the subject. Luminance of the patterns was adjusted so that they were clearly visible when presented for 10 msec in isolation without causing prominent after images. The visual angle subtended by a stimulus was about 2^{0}. The position on the screen of the S-patterns matched that of the corresponding part of the W-pattern. The room (and the display) was moderately tube-lighted. A response panel was provided with a starting key and three keys corresponding to three response alternatives.

Procedure. Fifty students of psychology participated and received a small payment. Each subject was tested individually. After a training series of 20 trials, 248 trials were administered in a random order, which took about 30 min. Trials were self-initiated. After the starting key was pressed, a fixation point appeared for 600 ms, followed by an empty screen for 300 ms. Then the two patterns (10 ms each) were presented, with one of four different SOAs: 20, 40, 60 or 100 ms. One second after the second stimulus the three S-patterns (A, B and C) belonging to the specific W-pattern, were displayed in a random order, from which the subject had to identify the subpattern which was presented in the flash together with the complete pattern. No feedback about the response was given. (In the training series, however, after an error the stimuli were shown once more but now for 1 s each). There were four conditions. In Condition W-S (96 trials: 4 SOAs × 8 Figure sets × 3 S-patterns) a pair S-W was presented; Condition S-W was identical but now with the S-pattern first and the W-pattern second. In Condition Sub-only (24 trials: 8 sets × 3 S-patterns) an S-pattern was shown followed by a grid of 9 × 9 lines which completely covered the S-pattern (SOA 20 ms). In Condition Complete-only (32 trials: 8 sets × 4 repetitions) a W-pattern was shown, so that none of the three alternatives was actually presented in the flash; subjects did not report noticing this.

Results

The percentages of correctly recognized S-patterns for Condition W-S and the control Condition S-W are presented in Figure 4a and 4b, pooled over figure sets (see Figure 3) and subjects. Each of the four base axes in Figure 4a and 4b represents the three S-patterns A, B and C, measured with four different SOAs.

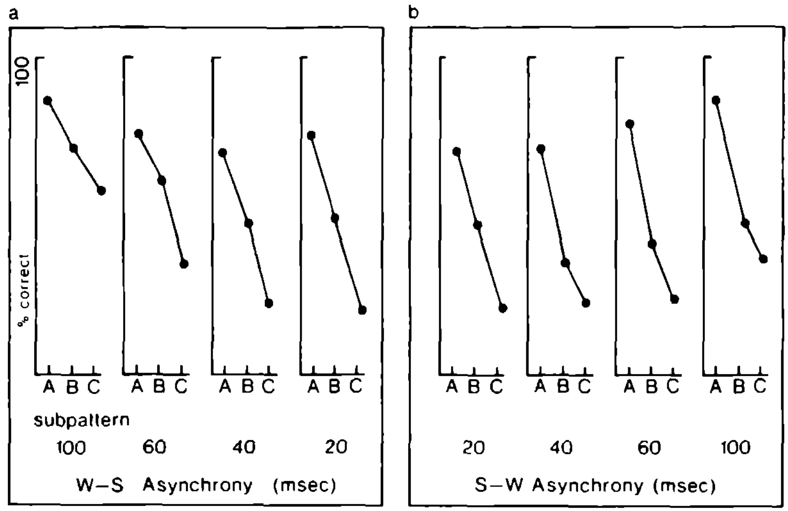


Figure 4. Percentages of correctly recognized subpatterns A, B, and C in Experiment 1. Panel a: Condition W-S, complete pattern first, subpattern second. Panel b: Condition S-W, reversed order of presentation.

The results are clearly dominated by an overall bias in favor of the better S-patterns (A > B > C). This bias cannot easily be explained by assuming a difference in detectability of the S-patterns as such: Although in condition Sub-only S-patterns of Type A seem correctly recognized somewhat more often than those of Type B and Type C (26.6%, 24.0%, and 24.2%, respectively), these scores do not differ significantly, $F(2,4)=0.469$, $p<.70$. In Figure 5a the "true" recognition scores are given after a bias correction has been applied (bias correction per figure set). The correction procedure is discussed below.

The corrected scores for the S-W order differ from those of the W-S order. An analysis of variance (ANOVA) was performed and the following two effects were significant. The recognition scores were better for larger SOAs, $F(3,9)=37.711$, $p<.001$. Order of presentation (W-S versus S-W) interacts with S-pattern, $F(2,6)=13.725$, $p<.01$. Further analyses showed that this interaction was mainly due to S-patterns A and B, $F(1,3)=29.442$, $p<.02$, while the interaction was not significant for B versus C, $F(1,3)=7.278$, $p<.08$, and for A versus C, $F(1,3)=6.702$, $p<.09$ (figure set as the error variable).

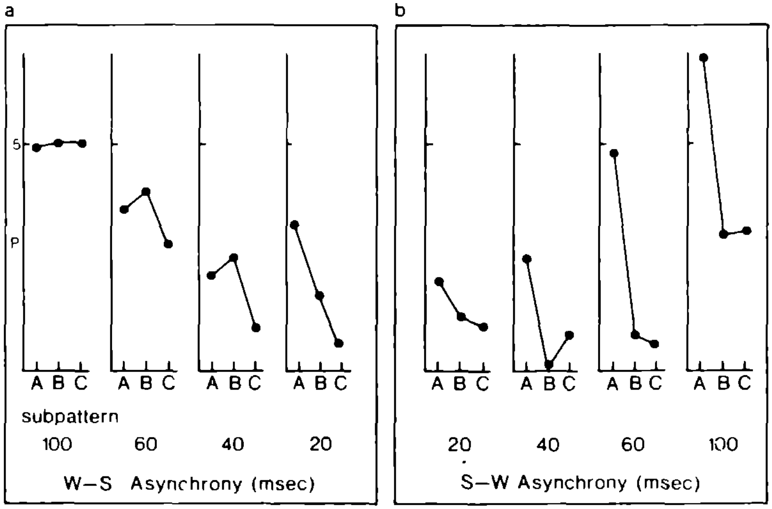


Figure 5. Proportion p of correctly recognized subpatterns in Experiment 1 after a response-bias correction. Panel a: Condition W-S (complete pattern first, subpattern second). Panel b: Condition S-W (subpattern first, complete pattern second).

We conclude that recognition of S-pattern B in the W-S condition is facilitated, as predicted on the basis of the assumption that the concurrent presence of the second-best interpretation of the complete pattern increases the visibility of the corresponding subpattern B.

Bias correction. At all SOAs, the A, B and C scores in Figure 4a and 4b reflect the intended preference of the corresponding interpretations: A > B > C. Table 1 shows that part of this ordering is due to a response bias: the errors indicate a bias towards the "better" S-patterns. The raw scores are separated into two components: a probability p_i of recognizing S-pattern (i) and a bias V_j for response alternative j. Hits (on the diagonal) are fitted according to the following expression:

$$p_i + (1 - p_i) \times (V_j)/(V_{sum}),$$

where V_{sum} is 1 + $V(2)$ + $V(3)$ (and i=j). Errors (off-diagonal scores) are fitted according to the second component only:

$$(1 - p_i) \times (V_j)/(V_{sum}),$$

(where $i \neq j$). The best fitting p and V values are determined by using an iterative chi-square minimalisation (program Minuit, Cern, Geneva). Thus we

Table 1
Uncorrected Scores in Experiment 1 for Conditions W-S and S-W

	Response category		
Subpattern	a	b	c
A	74	17	9
B	42	46	13
C	45	25	29

Note. Underlined values represent correctly recognized subpatterns. The scores are pooled over conditions W-S (W-pattern first, S-pattern second) and S-W (reversed order of presentation). Values are in percent.

have a response-bias correction revealing "true recognition" scores based on the overall error-distribution. Below we will give some arguments for the necessary independence between "recognizing" (p) and "not-recognizing" ($1 - p$).

Each subject may have a different bias distribution V_j for each of the four figure sets. Thus it would be optimal to calculate the *V*s for each of these 50 × 4 cases separately and to use these *V*s to correct the data per order of presentation (W-S and S-W) and SOA. However, the V estimation and the subsequent calculation of p need a considerable number of observations in order not to produce spuriously extreme values of V and p (corrections per subject and figure set yielded p-distributions that were bimodal and often extremely skewed). Only at the level of figure sets there were enough observations for the bias correction to yield stable results. First the V values were computed for each figure set from the responses on all SOAs and both orders of presentation. Then, with these *V*s the p values were computed on each level of SOA and order.

The correction requires independence between recognizing (p) and not-recognizing ($1 - p$). Three observations indicate such an independence. First, the results of Condition Complete-only are shown in Table 2 for the four figure sets together with the V_j values. These measures correlate strongly (r^2 =.91). In other words, the responses in Condition Complete-only as well as the errors in Condition W-S and S-W seem to be made only on seeing the W-pattern; not reporting the right S-pattern does seem to be based upon

Table 2
Scores in Condition Complete-only and Bias Values Computed for the W-S and S-W Conditions

	Response category					
	a		b		c	
Figure set	Score (%)	Bias	Score (%)	Bias	Score (%)	Bias
1	100	1	47	.68	40	.55
2	100	1	52	.52	7	.08
3	100	1	59	.69	29	.34
4	100	1	22	.25	21	.21

Note. For each figure set the results of condition Complete-only (a W-pattern without an S-pattern) is presented together with the computed bias values for the W-S and S-W conditions. Bias values are proportions of those of response category *a*.

being in a state of completely not-recognizing the target. Second, the error distribution was consistent over S-patterns. For each figure set and each order of presentation one error-score (a not-underlined cell in Table 1) was predicted using the error distribution in the other five cells. Although the correlation between predicted and observed values was weak (r^2 =.32), the averages did not differ, $t(7)=-0.61$, $p<0.56$. Third, at an SOA of 20 ms (both in W-S and S-W) the S-patterns and the W-pattern can be expected to fuse into one stimulus, with some line elements extra bright due to luminance summation. However, this does not seem to have resulted in differently perceived W-patterns for each specific S-pattern. The V_j-distribution for each of the four figure sets at SOA 20 ms is very similar to that of all four SOAs (r^2 =.82, $p<0.01$).

Effects of the previous trial. Trials were presented in a random order. Therefore, in about one out of every eight trials the stimuli were selected from the same figure set as in the immediately preceding trial. Frequently reported *sequential effects* on serial reaction time (e.g., Soetens, Boer, & Hueting, 1985) make it plausible that such a repetition of trials also will affect recognition scores. Applying the bias correction will tell more about sequential effects.

The S-pattern chosen on trial n - 1 (the previous trial) has a considerably increased chance to be selected on trial n, as can be seen in Table 3. (Analyses according to what S-pattern was actually presented on trial n - 1 yield

Table 3
Sequential Effects: Uncorrected Scores and Computed Bias Values

	Response on trial n					
	a		b		c	
Response on trial n-1	Score (%)	Bias	Score (%)	Bias	Score (%)	Bias
			Experiment 1			
a	53	1.0	30	0.7	17	0.2
b	39	1.0	45	1.5	16	0.4
c	39	1.0	27	0.9	34	1.4
			Experiment 2			
a	52	1.0	30	0.4	18	0.2
b	39	1.0	35	0.8	26	0.5
c	42	1.0	32	0.6	26	0.5

Note. The effect of the choice of subpattern chosen on trial n-1 on the choice on trial n for runs of trials with the same W-pattern. Uncorrected scores are given in percent; computed bias values as proportions of those of response category *a*.

similar but weaker effects and are not reported). However, the same response-bias correction procedure as described above (although in an extended version) makes it plausible that this sequential effect can be completely assigned to the response-selection phase. Comparison of Tables 3 and 4 shows that in case of the same figure sets the *V*s are increased, in contrast to the p values which appear unaffected on the whole (Table 4).

Table 4
Sequential Effects: Bias-Corrected Scores (p)

Response on trial *n*-1	Response on trial *n* a	b	c
	Experiment 1		
a	51	14	29
b	38	20	16
c	53	26	9
	Experiment 2		
a	41	36	27
b	46	49	42
c	56	56	43

Note. The effect of the choice of subpattern on trial *n*-1 on the choice on trial *n* for runs of trial with the same W-pattern. Values are in percent.

In the W-S order of presentation the complete pattern seems to activate the preferred A interpretation and the complementary B interpretation, facilitating the recognition of the corresponding subpatterns A and B. However, two problems have to be dealt with. The results for nearly simultaneous presentation in the previous experiment (SOA 20 ms) are somewhat unclear, both in Condition W-S and S-W. The gradual decrease from A to C seems to be a mixture of the data from the more extreme SOAs in both Conditions. Second, one can wonder whether the bias correction on the level of each figure set (F) instead of each combination of subject and figure set (S × F) did not distort the outcome. Therefore the experiment was replicated with the main purpose of collecting many repeated measurements.

EXPERIMENT 2

Method

Eleven subjects were run four times through a set of trials consisting only of figure sets 1 to 4 of Experiment 1 but each set was used two times. Thus eight replications were collected.

Results

The scores were corrected in two ways: once for each F as has been done in the previous analysis (Figure 6a and 6b, solid lines), and once for each com-

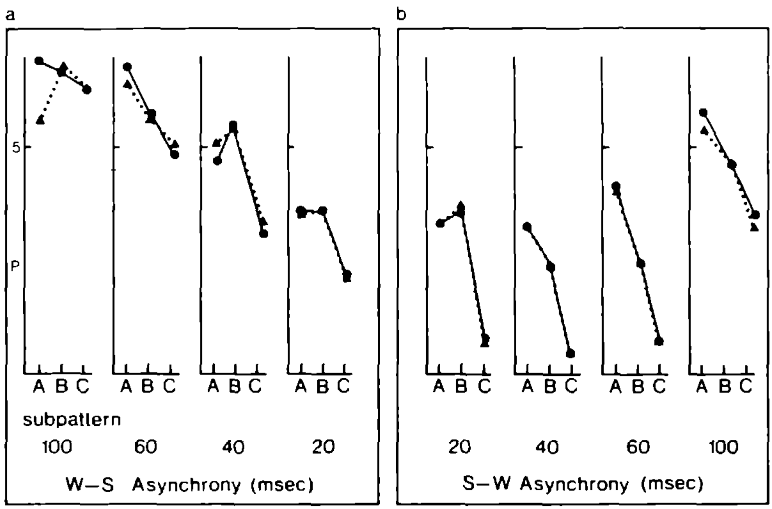


Figure 6. Proportion p of correctly recognized subpatterns in Experiment 2 after a response-bias correction. Panel a: Condition W-S (complete pattern first, subpattern second). Panel b: Condition S-W (subpattern first, complete pattern second). (Bias corrections on two levels of aggregation of the data are compared: Solid lines = bias correction for each combination of subject and figure set; broken lines = for each figure set pooled over subjects, as in Experiment 1.)

bination of S and F (interrupted lines). The close similarity between the two sets of corrected scores is an indication that the bias correction of the data of Experiment 1 for each F only is equivalent to a correction on the "optimal" level of S and F.

The ordering of scores over S-patterns for large SOAs in the S-W order is stable and similar to the response bias. Here, as in Experiment 1, another pattern emerges, but now already at 20 ms in the S-W order. ANOVAs were performed on the three larger SOAs of S-W (100, 60 and 40 ms; Group 1) on the one hand, and three successive SOAs (S-W = 20; W-S = 20 and 40 ms; Group 2) on the other hand. In general, the same effects appeared as in

Experiment 1. Larger SOAs yield better scores for all three contrasts (A vs. B, B vs. C, and A vs. C, $p<.01$, in each case) and group interacted with S-pattern only in the case of S-patterns A and B, $F(1,3)=14.963$, $p<.05$. No other effects were significant, which means that the interaction of group and S-pattern is stable over figure sets.

Runs of trials with stimuli selected from the same figure set now occurred twice as often as in the main experiment reported above. Again, the increased tendency to choose the same S-pattern in such a run can be ascribed to the bias component. No indications of facilitated "true" recognition are found (Tables 3 and 4).

An ANOVA on the raw data showed that practice interacted with only one other factor: The percentage of correctly recognized S-patterns A did not improve with practice as much as did B and C, $F(6,286)=3.42$, $p<.01$) perhaps partly because of a ceiling effect (A is recognized correctly in about 80% of the trials) and partly because subjects became aware of the less obvious response alternatives (B and C) and their assumed equal frequency.

Discussion

The response bias correction does not seem to be overly sensitive to differences between subjects. Its application on the results of Experiment 1 seems to have been correct. However, the correction is based on a *high threshold* model assuming an independence between recognition and (educated) guessing. High threshold models have been questioned on detailed second-guess data in a signal-detection setting (Coombs, Dawes and & , 1970; p. 189). However, we already argued that in the present study, the error responses were made on the basis of recognition of the W-pattern only. The independence of the bias and the recognition component is underlined by the analysis of the sequential data: A repetition of stimuli did not improve recognition, but it did increase the bias toward the same response category. A conventional signal-detection analysis, on the other hand, would not be optimal for these recognition data. To our knowledge, it requires a separate treatment of the hits and false alarms for each response category. However, the false alarm rate in one category depends on the bias for another category. These dependencies can be accounted for by combining the error distributions of all three response categories into one integral distribution, as has been done in the present analysis.

GENERAL DISCUSSION

We probed for the presence of interpretations of an unambiguous pattern (W) other than the preferred interpretation by looking at the recognizability of subpatterns (S) of that pattern presented immediately before or after it. The recognition scores, corrected for response biases, are presented in Figure 5 (Experiment 1) and Figure 6 (Experiment 2). Both sets of data show a facilitation of the recognition of the second-best (B) subpattern in the W-S condition relative to the S-W condition. That is, the S-W curves are more L-shaped than the W-S curves. We want to explain this interaction on the basis of the concurrent presence of the second-best interpretation (Option 2). In the introduction, we have already given arguments why we explain it neither on the basis of the relative strength of the alternative organizations nor by allowing for a probabilistic one-at-a-time dominance of the A or the B interpretation. We will now consider other alternative explanations of the data.

Masking

The paradigm of the present experiment is similar to that of studies directed at visual masking. Given the range of onset asynchronies which we used, numerous theories of masking predict an inhibiting effect of the complete pattern on the subpattern and vice versa. In fact, this is one of the reasons that the presence of (hidden) interpretations could only result in a facilitation of the recognition of compatible subpatterns *relative* to that of incompatible subpatterns. To explain the particularly good recognition of the B subpatterns in the W-S condition in terms of visual masking, we want to point out that aspects of a more or less primitive nature on which B subpatterns may differ from the other subpatterns are of primary importance. We want to argue that our material is balanced with respect to at least some low-level aspects generally assumed to play a role in masking. The subpatterns within each set were balanced with respect to the length of the constituting lines. In addition, the area enclosed by the smallest convex envelope was exactly the same within each set. Together, these two aspects made it safe to assume that brightness differences (Turvey, 1973) and differences in spatial frequency (Breitmeyer & Ganz, 1976) were negligible. It is our estimation that the same holds for "eccentricity", although there are small differences: Subpattern A of the left-most set in Figure 3, for example, is presented somewhat more parafoveally than, for instance, subpattern C of the same set; this may cause this particular A subpattern to be less susceptible to masking (Breitmeyer, 1980). Lateral inhibition or perhaps "metacontrast" may arise because of contrast gradients at points where lines appear or disappear from subpattern to complete pattern or vice versa (Lefton, 1973). Again, figure sets were sufficiently balanced in our estimation.

Features and Similarity

The subpatterns may have differed with respect to a host of other, more structural features. Hellige, Walsh, Lawrence, and Prasse (1979) established that from a SOA of 20 ms, a letterlike pattern was recognized better to the extent that it had less *featural similarity* with a mask presented after it. Without attempting to define featural similarity, we may assume that the similarity between a subpattern and complete pattern can be symbolized as: $s(W,A) > s(W,B) > s(W,C)$. Our overall results (better recognition of more similar subpatterns) clearly run against those of Hellige et al. This need not be a contradiction if Hellig et al. did not really measure featural similarity but simply the effect of the degree of overlap between target and mask. Overlap has been shown to attenuate target identification (Michaels & Turvey, 1979, Experiment E4). Instead of featural similarity, the term *template similarity* seems more appropriate with respect to their effects. Our complete patterns overlapped the subpatterns entirely, so the results of Hellige et al. do not seem to bear on the present results at all.

Nevertheless, the possibility cannot be excluded that our materials were not properly balanced with respect to some figural features. In fact, it is unavoidable that alternative organizations of a pattern differ more or less on aspects of a more local nature. We cannot discuss this problem thoroughly because the number of feature descriptions of even the simple patterns presently used is unlimited (Fu, 1972). However, we did inspect our materials with respect to this problem in two ways. First, what we assume to be the effect of a hidden pattern organization may have been the result of differences in similarity due to a relatively good match of essential features of the preferred organization and the B subpatterns which the C subpatterns lacked. We assessed the extent to which subpatterns were compatible with or violated the most plausible "essential features" of the preferred organizations: vertices, line continuations and part boundaries. B subpatterns are on the average only slightly more compatible than C subpatterns (33% vs. 27%, compared with, of course, 100% for the A subpatterns) and violate only marginally less of these essential features (52% vs. 61%; 0% for A). Second, subpatterns may differ irrespective of the relation with the preferred organization (note that the term *similarity* does not apply here). Most important, subpatterns B did not differ from A and C with respect to two so-called "primitive features" or "textons" (Julesz, 1981; Treisman & Gelade, 1980; Treisman, 1986): the number of line terminators and orientation. Only the amount of "closure" is somewhat less for B than for A and C (see, for instance, Figure 3, second set from the right). It is possible that this resulted in B subpatterns that are more vulnerable to masking (especially in the S-W condition). However, we want to remind the reader that the analysis of the results shown in Figure 6 did not reveal an interaction between figure sets and the differential effect for the B subpattern.

We want to conclude with the remark that the concept of similarity remains vague, all the more when one considers that the advantage for the concurrently present B subpattern can be said to be the result of its similarity to one of the (hidden) organizations of the complete pattern. We want to stress that the latter use of similarity relates to an aspect common to the perceptual organization of both complete pattern and subpattern, and not to a correspondence between local (geometric) features.

Masking Due to Incompatible Interpretations

In an earlier study (Leeuwenberg, Mens, & Calis, 1985), indications were found that even with SOAs as small as 30 ms, the first of two tachistoscopically presented patterns may be detected with difficulty when the organization of the second pattern cannot easily be expressed in terms of the first one. Such an effect of "incompatible" interpretations does seem to be in agreement with the declining recognition curves in the present condition S-W. It is what one expects on the basis of the relative preference for the three organizations. On the other hand, the W-S results deviate from such an explanation. Instead of preference, the concurrent presence of A and B interpretations is what we take to be the explanation.

Anticipation

One may suspect that subjects have learned to adapt to the task by anticipating the A and the B subpatterns before each presentation (either by conscious anticipation or in a more involuntary way). We have three arguments against such a learned anticipation. The first is that anticipation—at least partially—has been identified and accounted for. As stated before, subjects tended to choose the same subpattern as they did on the immediately preceding trial in the occasional instance that identical complete patterns were presented on subsequent trials (Table 3 and 4). This, however, proved to be a bias without an accompanying increase in the corrected scores. It may be assumed that knowledge of alternative subpatterns developing over a larger number of trials would share its workings with that of the short-term recall of the immediately preceding trial. If that is correct, long-term learning apparently affects the true, corrected recognition scores as little as does the immediate sequential effect. The second argument against an explanation in terms of learning is that an analysis of the raw scores of Experiment 2 showed that the recognition of both subpatterns B and C improved with practice more than that of subpattern A, instead of an improvement for B only. The third and last argument is that anticipation as such is not sufficient to explain the difference between the S-W and the W-S order of presentation.

There are aspects of the results which do indicate a sensitivity of the (bias-corrected) recognition to differences in the procedure of Experiments 1 and 2. The curves of the two experiments differ with respect to both the SOA at which the presence of the alternative interpretation is assumed to become effective, as well as the relatively good recognition of the B subpattern in the S-W condition of Experiment 2. We do not claim to be able to account for these differences, but we do want to make one suggestion. One distinctive aspect of the procedure of Experiment 2 was that we used four instead of eight figure sets as Experiment 1. This might have resulted in a less demanding task, leaving resources free for a more extended processing of the complete pattern. In that case, second-best interpretations may occasionally be preferred, which is similar to the relative freedom to shift from one perceptual organization to another under prolonged viewing.

Finally, a word about the relevance of the present results. Besides the tachistoscopic presentation, the nature of the material may be deemed artificial and leading to results of little importance. Indeed, the need to balance the subpatterns for as many local features as possible has led to the use of wirelike patterns that do not suggest surfaces, occluding edges, and such. Still, our complete patterns are equivalent to some natural shapes such as leafless trees and scratches on a hard surface. Furthermore, wirelike patterns are involved in numerous phenomena often used to demonstrate important aspects of perception (e.g., the Necker cube). One has also to consider the body of research using letters and digits.

We have presented an indication of the concurrent presence of a hidden shape interpretation. Although we think that such a hidden interpretation is more compatible with the view that perception optimizes the representation of structure than with the view that it minimizes processing costs, it is far from decisive. The present results do not tell us *how* the visual system selects these interpretations. The possibilities still vary widely. For instance, in an initial encoding phase, a very large set of interpretations may be generated; on the other hand, it could be that in general no superfluous interpretation is generated at all and that the system immediately converges on the small set of the preferred and (one) alternative interpretation(s). The latter would imply some sort of minimization after all. In any case, it is possible that we have shown only the outcome of shape encoding, not how this outcome is achieved.

REFERENCES

Attneave, F. (1982). Prægnanz and soap bubble systems: A theoretical exploration. In J. Beck (Ed.), *Organization and representation in perception.* Hillsdale, NJ: Erlbaum.

Ballard, D., Hinton, G., & Sejnowski, T. (1983). Parallel visual computation. *Nature*, *306*(*3*), 21-26.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, *83*, 1-36.

Breitmeyer, B. G. (1980). Unmasking visual masking: A look at the "why" behind the veil of the "how". *Psychological Review*, *87*, 52-69.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition*, *2*, 130-138.

Coombs, C., Dawes, R., & Tversky, A. (1970). *Mathematical Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Fu, K. S. (1972) On syntactic pattern recognition and stochastic languages. In S. Watanabe (Ed.), *Frontiers of patterns recognition* (pp. 113-137). New York: Academic Press.

Gottschaldt, K. (1929). Uber den Einfluss der Erfahrung auf die Wahrnehmung von Figuren, II [On the effect of experience on the perception of shapes]. *Psychologische Forschung, 12*, 1-87.

Gregory, R. L. & Gombrich, E. H. (1973). *Illusion in nature and art*. London: Duckworth.

Hatfield, G. C., & Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin, 97*, 155-186.

Hellige, J., Walsh, D., Lawrence, V., & Prasse, M. (1979). Figural relationship effects and mechanisms of visual masking. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 88-100.

Helmholtz, H. von (1962). *Treatise on physiological optics, I-III*. J. Southall (Ed. Trans. from the 3rd German edition of 1867). New York: Dover.

Herbart, J. (1850). *Lehrbuch zur Psychologie* [*Handbook of psychology*]. Leipzig: Leopold Voss Verlag.

Hochberg, J. & McAlister, (1953). A quantitative approach to figural goodness. *Journal of Experimental Psychology, 46*, 361-364.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290*, 91-97.

Kawamoto, A., & Anderson, J. A. (1985). A neural network model of multistable perception. *Acta Psychologica, 59*, 35-66.

Leeuwenberg, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology, 84*, 307-349.

Leeuwenberg, E., & Buffart, H. (1983). An outline of coding theory: a summary of related experiments. In H.-G. Geissler, H.F.J.M. Buffart, E.L.J. Leeuwenberg & V. Sarris (Eds.), *Modern issues in perception* (pp. 1-47). Amsterdam: North-Holland.

Leeuwenberg, E., Mens, L., & Calis, G. (1985). Knowledge within perception: Masking caused by incompatible interpretations. *Acta Psychologica, 59*, 91-102.

Lefton, L. A. (1973). Metacontrast: A review. *Perception & Psychophysics, 13*, 161-171.

McClelland, J., & Miller, J. (1979). Structural factors in figure perception. *Perception and Psychophysics, 26*, 221-229.

Michaels, C. F. & Turvey, M. T. (1979). Central sources of visual masking: Indexing structures supporting seeing at a single brief glance. *Psychological Research, 41*, 1-61.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226-254.

Orbach, J., Ehrlich, D., & Heath, H. A. (1963). Reversibility of the Necker Cube: I. An examination of the concept of "satiation of orientation". *Perceptual and Motor Skills, 17*, 439-458.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9*, 441-474.

Reed, S. K. & Johnsen, J. A. (1975). Detection of parts in patterns and images. *Memory & Cognition, 3*, 569-575.

Perkins, D. (1983). Why the human perceiver is a bad machine. In J. Beck, B. Hope & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 341-364). New York: Academic Press.

Sadler, T., & Mefferd., R. (1970). Fluctuations of perceptual organization and orientation: stochastic (random) or steady state (satiation)? *Perceptual and Motor Skills, 31*, 739-749.

Simpson, G. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*, 316-340.

Soetens, E., Boer, L., & Hueting, J. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 598-616.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behaviour, 18*, 427-440.

Treisman, A, & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80*, 35-52.

Van Tuyl, H. (1980). Perceptual interpretation of complex line patterns. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 197-221.

Weisstein, N., & Maguire, W. (1978). Computing the next step: psychophysical measures of representation and interpretation. In A. Hanson & E. Riseman (Eds.), *Computer Vision Systems* (pp. 243-260). New York: Academic Press.

Appendix

Table A1

The complexity of components of the A, B, and C subdivisions of the four figure sets (Experiments 1 and 2) is calculated on the basis of the coding rules proposed by Leeuwenberg (1971). Letters refer to components depicted in Figure 3.

Component	Code	Reduced code	Information load
a	qmam	→ qS(m,a)	4
b	qlam		4
c	qmambm	→ q<m><ab>m	5
d	qm		2
e	qm[am]m	→ qS(m,[am])	5
f	qm[amam]m	→ qS(m,[2(am)])	6
g	qmamam	→ q2(ma)m	5
h	qlamal	→ qS(la,m)	5
i	qmambmbmam	→ q(S(<m><ab>,m)	6
j	qmamambmbm	→ q<m><2(a)2(b)>m	7
k	qamalbmlamalm	→ q<(S(la,m))><(bm)(m)>	8
l	qlamam[bm]p	→ ql2(am)[mb]p	8

Note: q=position and orientation; 2(*x*)=*xx*; <*x*><*yz*≥*xyxz*; []=bifurcation operator; →=reduction of left code into right code; .=parameter contributing to the load.

Received June 10, 1986
Revision received September 3, 1987
Accepted December 15, 1987 ◆

The Perceptual Representation of Array-Element Position: Grouping and the Span of Apprehension

Lucas H. M. Mens and Emanuel L. J. Leeuwenberg

University of Nijmegen, the Netherlands

When looking at an array of identical and regularly spaced elements (e.g. a grating), one becomes aware of a severe limitation in the number of elements that can simultaneously be attended to, testifying a limited *span of apprehension*. A small number of elements allows for a representation of each individual element; many elements yield a texture-like impression. A coding model is presented as a tentative explanation of this limited span of apprehension. In the model, the relative position of pairs of elements is completely specified with binary ordering relations. Only a limited possibility of combining binary relations on one hierarchical level is assumed. The model predicts that the "individuated" representation of 4 to 6 elements requires a representation of deeper hierarchy than one of 1 to 3 elements, but less deep that still larger arrays. In a same/different task, subjects indicated whether two successively presented arrays of identical and equidistant shapes contain an equal number of those shapes. As predicted, the reaction times increased sharply between 3 and 4. Some support for the predicted discontinuity between 6 and 7 is also presented. Instead of indicating a fixed capacity, the limited range of processes such as *subitizing* and *focal* perception may well be explained by structural constraints—such as the ones imposed by the present model—on the format of representation.

GENERAL INTRODUCTION

This paper is about the representation of positional information in a visual array. A tentative model will be presented in which the ordering relations between the individual elements of a linear array are *explicitly* represented in

---

A preliminary draft of this paper was presented at the eight European Conference on Visual Perception in Peniscola, Spain, 1985. An abstract can be found in Perception, 1985, *14*, A10.

This research was supported in part by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

The authors wish to thank Noud van Kruysbergen for his assistance in these experiments.

a propositional code. Constraints will be formulated which lead to the prediction of a *discontinuous* increase in representational complexity with the number of array elements, modelling the discontinuous increase in perceived complexity.

In his classical study on grouping effects in visually perceived forms, Goldmeier (1936/1972) notes the fundamentally different appearance of Fig-

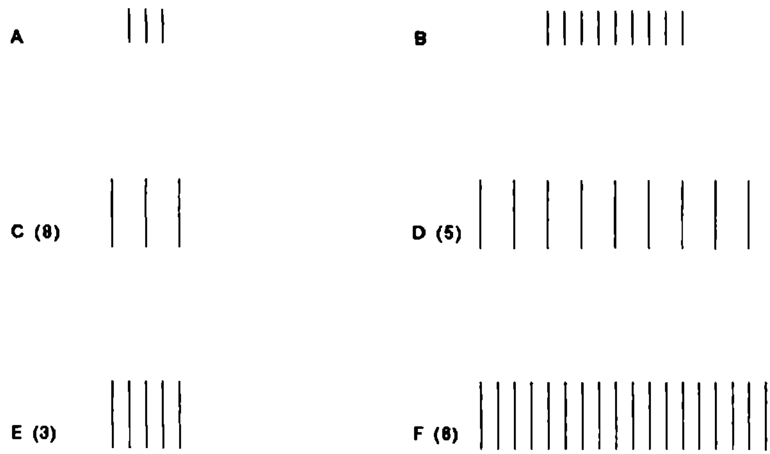


Figure 1. The effect of the number of elements on number-preserving similarity judgements. Eight of 11 subjects judged **A** to be more similar to **C** than to **E**; only 5 judged **B** to be more similar to **D** than to **F**. Similar data from Goldstein (1972/1936, p.48) are: 7.5 out of 10 for **A,C** and 3 out of 12 for **B,D.**

ure *1B* as opposed to Figure *1A*. In Figure *1A*, each element seems to be a unique shape in its own right, while the elements in *1B* are only part of the "material" of the global percept. Accordingly, subjects judge *A* as more similar to *C* than to *E*, while *B* is judged to be more similar to *F* than to *D*. Looking at an element in *B* gives one the impression that the visual system cannot establish the position of each element unequivocally. As a result, it is difficult to discern all elements of the array simultaneously, which results in a texture-like appearance. In contrast, *1A* seems to allow for an "individuated" perceptual representation of each element.

One may wonder whether the number of elements is the crucial difference between *A* and *B*. It certainly is not the *only* factor. Kimchi and Palmer

(1982) emphasize that decreasing the absolute size of the elements may induce a shift towards texture-like appearances. One can observe a possibly similar phenomenon by decreasing the brightness of a random-dot display to a very low level. This leads to an increase of perceived numerosity. The common cause may be a decrease in the sensory signal/noise ratio. Thus, absolute size seems to be a factor but probably only in nearly-threshold cases. Another possible factor is the size of the elements *relative* to the size of the array as a whole (compare Figure *2A* and *2B*). In a pilot study we established that the relative size had only a very small effect on the number of number-preserving similarity judgements. Therefore, the number of elements as such does seem to be a crucial factor, within the limits of sensory sensitivity. Of course the limited capacity to discern perceptual elements is directly linked to the old concept of the *span of apprehension* (see Craik & Levy, 1976) and to later operationalisations which have led, for instance, to the *magical number*(*s*) of STM. However, the model which we want to propose is intended for visual perception, and not for the rehearsal of verbal material, so we will not discuss the various theories of item association within a STM list (see McNicol & Heatcote, 1986). We will merely quote Shiffrin (1976) who states that "order information may be more basic than item information in short-term retention" (p. 203) and that "the advantages of chunking and spacing, and the improved retention of positional information may result from active processing on the part of the subject.." (p. 204). It is exactly the need for active processing of positional information which we think is basic to the problem of perceptually keeping many elements apart from each other. We propose the chunking of elements into a hierarchical code as the only way of coping with larger numbers of elements.

Another indication of the limited ability to retain positional information can be found in a study of Attneave and Curlee (1983) who asked subjects to imagine a point wandering in a matrix according to spoken instructions. After twelve of these instructions, the final position in (imagined) matrices of size 3 *3 was correctly indicated in about 90% of the trials, whereas matrices of size 4 *4 and larger proved to be considerably more difficult. Performance on the larger matrices improved when the subjects were instructed to use a (hierarchical) structuring into smaller sub-matrices. The authors argue for the equivalence of imagined and perceptual positional information; a pixel-like representation is dismissed.

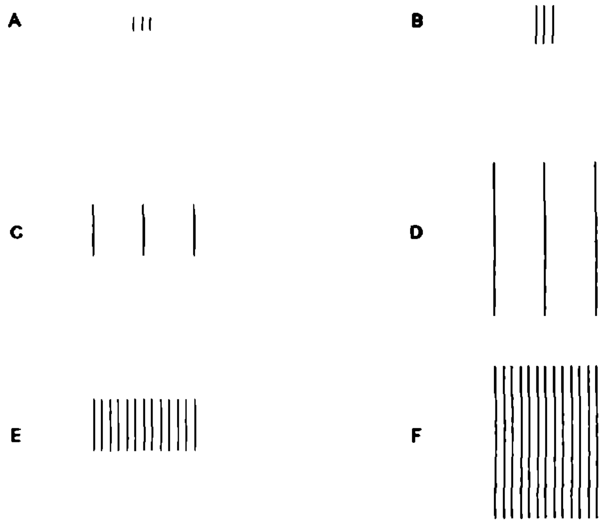


Figure 2. Six arrays were constructed using varying number (3 versus 7) and lengths of lines (1, 2 or 3 units, compare 2*A* and 2*B*). Two arrays were derived from each array **A** and **B**, both three times as large: arrays *C* and *D* are simple enlargements and therefore preserve the number of elements; in *E* and *F* the spacing between the lines is the same as in **A** and **B**, keeping the (local) texture constant. Subjects had to indicate whether *C* (*D*) or *E* (*F*) was more similar to *A* (*B*). The number of elements in **A** and **B** had a strong effect on the number-preserving responses (that is, *C* and *D*), but the length of the elements in **A** versus **B**, and therefore the relative size of the elements, did not.

Also indicating a limited span of apprehension are studies on number perception. Here too, small numbers of elements are processed in another way than large numbers. Reaction times of a naming task for displays within the range of 1 to about 3, 4 or 5 elements (depending on mode of presentation and on the analysis) are shorter than those for larger arrays. The seemingly "immediate" perception of the numerosity of small sets of elements has led to the coining of the term *subitizing*. Several authors have made an attempt at explaining subitizing. Following Woodworth and Schlosberg (1954), in some studies fast subitizing responses are predicted for overlearned configurations of elements (e.g. 4 is a quartet; Mandler & Shebo, 1982; Simons & Langheinrich, 1982). Others have proposed a bottleneck due to some limited capacity, be it in the number of receptive fields acting as "counting units" (Atkinson, Campbell & Francis, 1976), or in the number of slots in short-term memory to keep track of which items have been counted (Averbach, 1963). It has also been claimed that subitizing is but one manifestation of a general process of discrimination, special only because with small numbers of elements, the uncertainty of discrimination does not exceed a critical threshold (Woodworth & Schlosberg, 1954; van Oeffelen & Vos, 1982). The actual value of this threshold is empirically established. Finally, Klahr (1976) has to be mentioned who modelled subitizing in an algorithm which attempts template matches of increasing size until a template big enough for the stimulus is found. Then the number response is derived from an iteration index. The maximal size of the template is a free parameter, estimated on the basis of subitizing data.

What all these explanations do not give is a principled account for the range of arrays which can be subitized; a theory about representation and associated processing is lacking. Below we will propose a model about limitations on the representation of positional information which may partly explain subitizing. It is only a partial explanation because with respect to subitizing it is important to distinguish between the initial phase of the perceptual encoding and retention of the pattern, and the subsequent phase of number retrieval. These processes probably have distinct functional and developmental bases, as is indicated by the following two facts: young children do not subitize, but instead start with counting even very small arrays (Gelman & Gallistel, 1978); on the other hand, even infants of age 2 show dishabituation if an array of 2 changes into one of 3 but not if larger arrays are used (Starkey & Cooper, 1980; Strauss & Curtis, 1981; Antell & Keating, 1983). Therefore, an early form of the "immediate apprehension" of a small number of items does seem to exist, independent of the ability to derive a numeric label from this perceptual knowledge. We want to focus on the perceptual process, not on the arithmetical and verbal second one.

We will restrict ourselves to the "immediate apprehension" of horizontal arrays of identical vertical lines. The advantage of such linear arrays is that the global outline qualitatively stays the same with varying numbers of elements. The reason for the restriction to identical elements is that they can be kept apart solely on the basis of their individual position, as they do not differ in (meaningful) content. More specifically, we assume that the preferred shape representation of the presently used materials is independent of the number of elements (it is, roughly speaking always "an array of identical elements"), and that the task demands a second representation which does preserve the numerosity. In this respect, it is convenient to think of this second representation as being *complementary*, supplying information relevant to the task lacking in the first representation (Collard & Buffart, 1983). Note that "an array of identical elements" also seems to be an appropriate characterisation of the texture-like interpretation. This implies that we assume that the two interpretations (the individuated one discerning each element and the texture-like one) differ only in the presence of the complementary representation which allows for a more *complete* representation.

In the following, we propose a model for an *individuated* representation which is complete with respect to the ordering relations. We will predict the processing load for a given array-length from the model. At this point we want to emphasize that, although one may instruct subjects to be accurate, one cannot prevent that a texture-like representation is used which is incomplete but which offers an "easy way out", especially with very large and/or briefly presented arrays. We will also see that conditions differ in discriminability on the basis of such a texture-like representation.

A tentative model of positional information in a linear array

Our aim is to model the representation of a series of elements in such a way that information about the number of elements is preserved. In the introduction, we have already dismissed a coding in terms of spatial coordinates. Another possible solution would be to assume an early counting process generating explicite numbers as output. Pattern "I I", for example, can be coded as "2 *(I)". The problem then is how to distinguish "2" from another number, say "4". We see two options. The first is to represent the numbers in an analogue dimension, which re-introduces the shortcoming of a discrimination account discussed above and which, implicating a *dense symbol system* (Pylyshyn, 1984, p. 199), is not meaningful in a cognitive view of mind. The second way to distinguish numbers is by reference to knowledge of arithmetical rules (e.g. "4=2+2"), which we consider to be inaccessible to perception. Instead of using numbers, we will encode geometric information in a discrete symbol scheme using a binary proposition which may be thought of as: *-is-left-of-* (cf. Sutherland, 1968; Foster, 1985). For convenience, the symbol "$<$" will be used to represent this relation. Using the notation a for the first element and b for the second, sequence "| |" can be coded as "$(a<b)$". Both a and b are to be understood as place tokens expressing the presence of an (unspecified) perceived object at a certain position a and b respectively. We require the code to be complete. For instance, code "$(a<b),(a<c)$" of a sequence of three elements is incomplete because it does not specify the ordering with respect to b and c. On the other hand, code "$(a<b),(b<c)$" is complete. It may be noted that the latter code implies a superordinate ordering of the propositions "$(a<b)$" and "$(b<c)$". In contrast to coding models such as that of Foster (1985), we require such an ordering to be expressible using the same *binary* relation. In order to be able to do so, we introduce the notion of a chunk: each relation (e.g. "$(a<b)$") may enter a

$(a<b)\ (b<c)$ ← Chunk Level 1
$(x < y)$ ← Chunk Level 2 (redundant)

superordinate level as one chunk x. Hence: An assumption is that the expressions on the first Level (that is: $(a<b)$ $(b<c)$) already satisfy the completeness requirement. The reason for this assumption is that b in both instances denotes one and the same place token of a perceived object. As a result, the superordinate expression "$(x<y)$" is considered to be redundant. It

$(a<b)$ Chunk Level 1
$(x < c)$ Chunk Level 2

clearly is not redundant in the following code of three elements:
Expression "$(x<c)$" does contain new information and therefore adds an independent level. In Table 1a representative selection is given of alternative codes for sequences of length 2 to 8. Two possible measures of complexity are supplied in Table 1. The number of binary relations increases linearly with the number of elements, which of course is not informative. Moreover, the number of binary relations does not distinguish between alternative codes for a given number of elements. The number of *Chunk Levels*, on the other hand, displays discontinuities between 3 and 4 and between 6 and 7. Thus we will use Chunk Level as a predictor for the complexity of an array.

As an alternative to the notation in Table 1, brackets may be used to indicate the hierarchical nesting of levels in perhaps a more instructive way. A code of a sequence of 8 may be written as follows:

$$(\ ((a<b)<(c<d)) < ((e<f)<(g<h))\ )$$

Table 1
Alternative codes for arrays of length 2 to 8

Length	Code	Number of relations	Number of levels	Preferred
2	(a<b)	1	1	*
3	(a<b) (b<c)			
	x y	2	1	*
	(a<b)			
	(x < c)	2	2	
4	(a<b) (c<d)			
	(x < y)	3	2	
	(a<b) (b<c)			
	(x < d)	3	2	
5	(a<b) (b<c) (d<e)			
	(x < y)	4	2	*
	(a<b) (c<d)			
	(x < y)			
	(z < e)	4	3	
6	(a<b) (b<c) (d<e) (e<f)			
	(x < y)	5	2	*
	(a<b) (c<d)			
	(x < y) (e<f)			
	(p < q)	5	3	
7	(a<b) (b<c) (d<e) (e<f)			
	(x < y)			
	(s < g)	6	3	
	(a<b) (c<d)			
	(x < y) (e<f) (f<g)			
	(s < t)	6	3	
8	(a<b) (c<d) (e<f) (g<h)			
	(x < y) (p < q)			
	(s < t)	7	3	

Note. A representative selection of codes for arrays of 1 to 8 elements. The place tokens are given the letters <a,b,..,h>; Chunks are referred to by the letters <p,q,..z>. A selection between alternative codes (identified with an asterisk) is made on the basis of the number of non-redundant Chunk Levels.

Summing up, we have proposed a representation of arrays of identical elements with the following properties:

(1) A code must preserve the information about the ordering completely.
(2) Place tokens are used linking the array elements to the code.
(3) The relative position of a pair of elements is coded in a binary ordering relation defined over their place tokens: "(*a*-is-left-of-*b*)".
(4) For the complete coding it may be necessary to express the ordering relation between ordered pairs on a higher level (again using binary ordering relations). This amounts to the formation of *Chunk Levels.*
(5) Identical place tokens in two ordered pairs bound in a (higher order) chunk render that chunk redundant.
(6) The number of Levels with non-redundant chunks is a measure for the complexity of the code.

EXPERIMENT

The model presented above implies that arrays of length 1, 2 and 3 are represented without the formation of a second Chunk Level. Arrays of length 4, 5 and 6 are represented with two Chunk Levels. Therefore, an extra load is predicted for arrays with more than three elements. To show this difference, the following task is chosen. An array (S1) made of a number (n) of identical line patterns is presented for 1 sec. After that, another array (S2) is shown containing the same number of elements (n) or a different number of elements (one less: n-1, or one more: n+1). The second array is shown for only 150 msec to prevent the effect of subsequent eye-fixations. Subjects have to respond as fast as possible with a *same* or *different* response.

A marked increase in RT of *same* trials is predicted for arrays with more than three elements in S2. In fact, the same increase is indicated by the model between 6 and 7 elements,[1] but it is likely that increases at higher numbers will be obscured by another effect: as is demonstrated in Figures 1 and 2, larger arrays elicit an increasing tendency towards a texture-like representation which is not as accurate as the representation proposed above (as far as the number of elements is concerned) but which allows for a fast same/different response.

[1] Indeed, an infinite sequence of transitions to higher levels follows from the model after 3, 6, 12, 24, .. elements, characterised by: $3 * 2^n$ (n=0,1,2,..).

We expect that the option of a texture-like representation also will obscure the results of *different* trials because length and/or density differences between two arrays of unequal number are more informative. Using random dot displays, Foster (1978) concluded that "Different number" judgements are mediated (in part) by a shape-independent representation consisting of area and density features. "Same number" judgements, on the other hand, were found to be based on a representation of the shape of the pattern. Therefore, if the discontinuity is present between 3 and 4 elements in S2 for the *different* trials, we predict it to be less distinct than for the *same* trials.

Method

Materials

Two series of patterns consisting of horizontal rows of line patterns (circles, triangles, squares or crosses) were presented on each trial. Series 1 (S1) contained 1 to 8 elements, Series 2 (S2) contained the same number, one less or one more (1 to 9). S2 was always constructed using a different shape as for S1. These series were computer generated on a display (Vector General Graphics Display Series 3, Model 2D3 with P4 phosphor) positioned at a distance of 250 cm in front of the Subject. The visual angle subtended by each element was 0.12 degrees. The spacing between elements varied randomly from trial to trial. The maximal visual angle of the largest series was 4.4 degrees. Furthermore, the eccentricity of each series was varied at random with a minimum of 0.4 and a maximum of 1.2 degrees visual angle in a randomly chosen direction from a fixation point. All these (random) variations were added in order to avoid responses on the basis of the perception of apparent movement, the length of the array, or other (in)variant aspects not under study. The experimental room was dimly lit. A response panel was provided with a starting key and separate keys for *same* and *different*.

Procedure

Fifteen students of psychology participated to fulfill introductory course requirements. Each subject was tested individually. After a training session of 20 trials, 368 trials followed in a random order which took about 40 minutes. Trials were self-initiated. After the starting key was pressed, a fixation point was shown which preceded S1 by 600 msec. S1 was presented for 1000 msec followed by one of four different intervals (ISI): 0, 8, 30 or 400 msec. Finally S2 was presented for 150 msec. A trial was completed when the Subject selected the *same* or *different* key. Reaction times were measured from the onset of S2. Instructions emphasized both speed and accuracy. In case of a RT beyond 900 msec the text "Please respond faster" was displayed. Wrong responses were signaled to the Subject with a buzz. Each combination of the numbers of elements in S1 and in S2 (3 *8 - 1, no sequence of 1 and 0 elements of course) was presented at each ISI 16 times (using each pattern ("circle" etc.) 4 times in S1).

Results

The average RT's and the percentages correct for the three Conditions (*one less*, *same number*, *one more*) are shown in Figure 3. An ANOVA was performed for each of the three Conditions on the median RT's of each Subject per ISI. (A separate analysis on the correct responses yielded very similar results and will not be reported.) A significant increase in RT with increasing number of elements in S2 is present in all three Conditions (*one less*: $F(6,392)=4.57$, $p<0.01$; *same number*: $F(7,448)=6.27$, $p<0.01$; *one more*: $F(7,448)=4.60$, $p<0.01$). ISI has no significant main effect, nor does it interact with the number of elements in S2. Significant differences between RT's for each number of elements in S2 per Condition as indicated by a Student Newman-Keuls test ($\alpha=0.05$) are depicted in Figure 3 by the dotted lines. The predicted jump in RT from 3 to 4 elements in S2 is evident.

The error rates covary with latencies, suggesting that the RT results are not caused by speed-accuracy trade-offs. As a whole, the errors increase gradually with the number of elements. There are no significant differences (SNK, $\alpha=0.05$) between two subsequent array-lengths, except for S2: 5→6 and S2: 6→7 in Condition *one less*.

Discussion

A steep increase of reaction times between 3 and 4 elements in S2 is present in Condition *same number*, as predicted. It is, however, not present in the other two Conditions. As stated above, the lack of a clear effect in *different* trials can easily be understood as the result of the discriminability of S1 and S2 in these trials due to a "simple" difference in area and density. Then, a

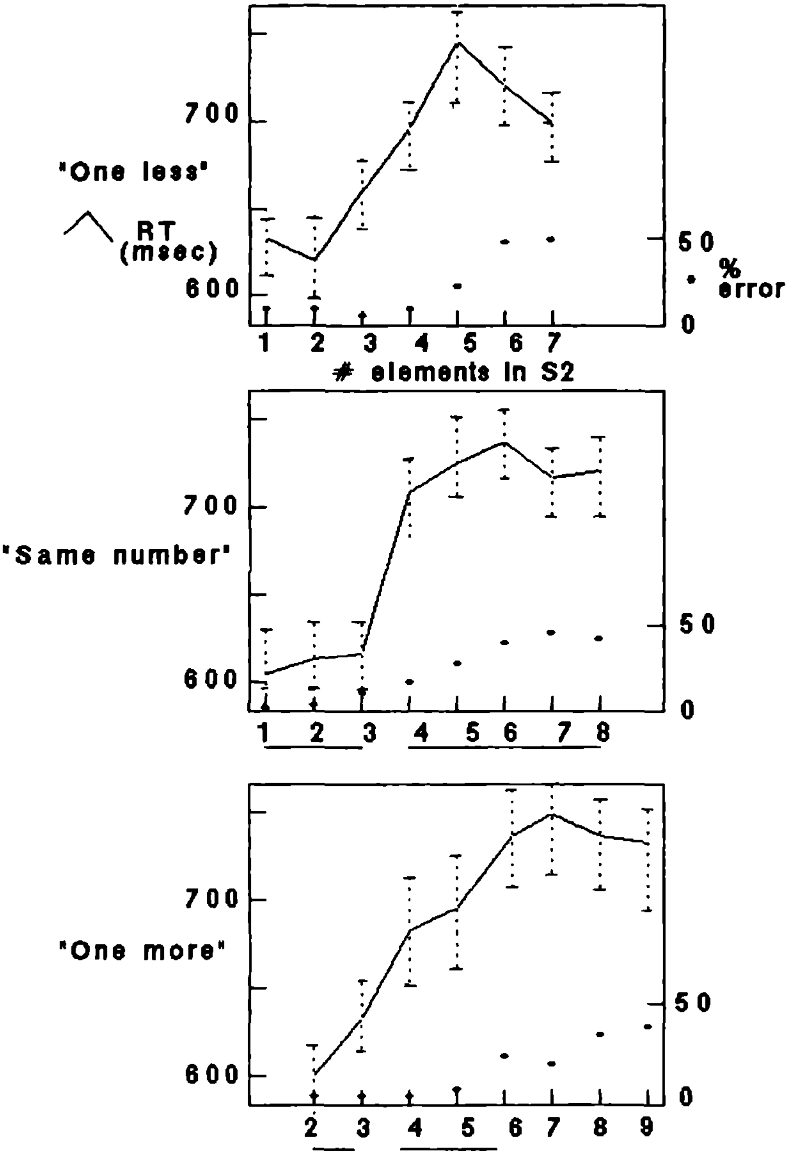


Figure 3. *Same/different* judgements times (in msec) and percentages error. The second array (S2) contains one element less, the same number or one element more than the first array (S1). Subsequent RT's of a curve not connected by one of the horizontal lines at the bottom (see *same number*) differ sifnificantly (SNK grouping, α=0.05).

complete, "individuated" representation is not the only basis for the same/different response.

The predicted extra increase in complexity from 6 to 7 elements could not be established. Instead, subjects seemed to make a fast but inaccurate response based on a less exhaustive representation of the largest arrays. In a subsequent experiment, we gave one of our subjects extensive practice on the same task (4 replications). This subject was selected because he produced the shortest mean latencies. The subject was pressed to make no errors, even on the larger arrays. In order to discourage errors on the larger arrays, only arrays of length 4 to 8 were administered. RT's decreased with practice ($F(3,15)=5.59$, $p<0.01$) but practice did not interact with the number of elements in S2, nor with the difference between the number of elements of S1 and S2. The effect of array-length is not straightforward. See Figure 4 for

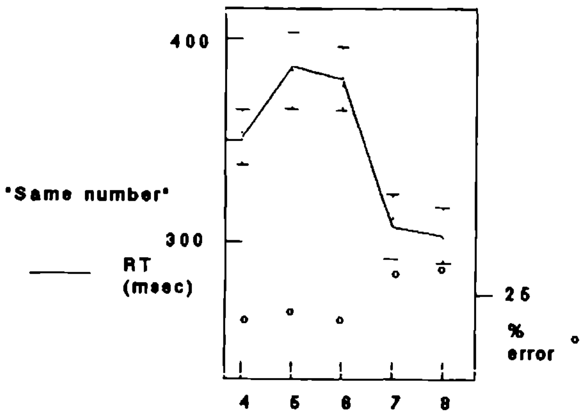


Figure 4. *Same/different* judgement times and percentages error of subject C.S. (four replications of the same task) for the condition *same number.*

the results of the crucial Condition *same number.* The number of errors increases considerably between 6 and 7, but inspection of Figure 4 reveals a large drop in the RT's at the same point. Although the RT results run against our hypothesis, we are inclined to interpret the large speed-accuracy trade-off as a support for the predicted increase in complexity of the code. As argued before, subjects may rely on an inaccurate texture representation if it leads to a fast response, but the same may well happen if not enough effort or inspection time is available to reach an individuated representation. In spite of his considerable skill, this subject seemed to be able to reach that individuated representation for arrays beyond length 6 much less easily than for smaller arrays, in agreement with the predicted increase in complexity.

It has to be noted that the 6-7 effect observed in the single subject study seems to have a qualitatively different basis than the 3-4 effect observed in the main experiment. RT's and errors in the main experiment covary with the number of elements in the array presented in the second position (S2, see Fig. 3). The responses of the single subject were clearly dominated by the first array (S1) instead, as an inspection of the data showes. Furthermore, of the three ISI's (1, 8, 30 and 400 msec) the longest ISI yielded increased RT's (SNK grouping, α=.05) while ISI did not have any effect in the main experiment. This strongly suggests that the single subject entertained strong expectations about the second array based on the perception of the first array, not only of the arrival time but also of the difficulty of gaining a sufficiently good impression. Arrays of length 7 and 8 seem to be processed in a much more superficial way.

Thus, even a proficient subject who received considerable practice displayed considerable difficulty with arrays of length 7 and larger. We conclude that the span of apprehension does not shift continuously with practice. Firm support for the predicted transition to a higher-level representation between 6 and 7, however, is lacking and seems to require an even easier task than the present same/different judgement about two successively presented arrays. One extra piece of evidence can be found in Goldmeier's (1936/1972) similarity judgment studies. Goldmeier's results were gathered under prolonged viewing conditions. It is obvious that similarity judgements and individual scanning strategies will not result in clear-cut results. However, using triads such as in Figure 1, number constancy (I→II in Figure 1) was preferred for up to about 6 elements. At 7 and 8 elements, density-preserving judgments appear, which dominate from 9 upwards. Thus we find at least some extra support for the predicted increase in complexity beyond 6.

GENERAL DISCUSSION

We conclude that the increase in complexity of the codes for arrays of length 4 and larger was confirmed by the increased time taken to decide whether two subsequently presented arrays contain an equal number of elements. Indications for the predicted increase in complexity from 6 to 7 were found in a marked speed-accuracy trade-off at that point. Furthermore, we believe that larger arrays cán be perceived without excessive inaccuracy, but only on the basis of a complex representation constructed after prolonged viewing. Instead of being explained by a fixed capacity limitation ("the number of slots"), the difficulty which perceivers experience to keep identical elements apart (the "span of apprehension") is modelled as an essential limitation on the structure of the mental representation.

Of the six properties of the model listed above, at least three of them (properties 3, 4 and 5) are straightforward assumptions while the others can be considered to be generally acceptable statements. It is clear that the present results support the model as a whole, but hardly each individual assumption. It is for this reason that the model still has to be considered as a tentative one, merely intended as a starting point in understanding the phenomena at hand. One clear shortcoming in this respect is that it is far from evident how such a coding scheme would apply to nonlinear patterns.

The present results point at limitations in dealing with merely a handful of visual items. In contrast, early retino- and spatiotopic feature maps exhibit a vast spatial capacity, in line with the—disputed—massive visual store: the *icon*. But of course, we did not use a feature detection task. As Treisman and Paterson (1984) argue, an early parallel and preattentive stage of encoding may suffice for feature detection. Combining features into a perceptual object, on the other hand, may require focused attention in order to coordinate the feature maps involved. Focused attention is required to an even greater extent for the present task (and for subitizing), where the elements possess no distinctive (emergent) feature whatsoever. We feel that the "individuated" representation of each element, needed to be able to respond to the exact number, amounts to a singular task for the visual system.

If perception is limited in this respect, what about the introspective density of imagery? We already referred to the facilitation of subdividing the imagined space when keeping track of a moving spot, which was experienced by subjects of Attneave and Curlee (1983). It may be the feeling of knowing what cán be imagined, instead of the image itself, which causes the introspective density. Simons and Langheinrich (1982) showed that perception of the number of elements of an array is facilitated when the local parts form a simple or overlearned configuration. It seems that a "hierarchical access-structure" is the central concept. In other words, a structured context may help to circumvent the severe limits of what can be kept apart within one "act of apprehension".

In normal situations the number of many identical items and the ordering between them is unstable and not of vital importance. Nevertheless, we feel that the problem of focusing attention to a number of items simultaneously may be a very general one, shaping everyday behaviour in many instances. For example, a great number of sequences is grouped in a very small number of elements, mostly in triplets, be it action sequences ("I count to 3", ball/floor contacts in sports) or conceptual sequences (three-colored traffic lights, *upper-middle-lower*, *past-present-future* etc). An ergonomic implication of the present results is that one should be aware of severe perceptual limitations in cases when items differ in only one respect. Crudely stated, performance may deteriorate when more than three identical items have to be kept apart.

In some respects, the present model is equivalent to one proposed by Wickelgren in 1967 which was brought to our notice by an attentive reader of this manuscript. Wickelgren explained the (hierarchical) grouping in triplets shown by subjects in a short-term digit rehearsel task (also observed by Ryan, 1969; Kahneman & Henik, 1977) with the assumption that only three place markers are available (*begin-middle-end*). This seems to be a straightforward ad-hoc assumption. Nevertheless, it needs to be clarified whether the major assumptions underlying our linear ordering model have an independent justification. Some possibilities are infant perception studies and micro-genetic alterations of the arrays.

REFERENCES

Antell, S., & Keating, D. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*, 695-702.

Atkinson, J., Campbell, F., & Francis, M. (1976). The magic number 4 ± 0: a new look at visual numerosity judgements. *Perception*, *5*, 335-342.

Attneave, F., & Curlee, T. (1983). Locational representation in imagery: A moving spot task. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 20-30.

Averbach, E. (1963). The span of apprehension as a function of exposure duration. *Journal of Verbal Learning and Verbal Behaviour*, 2, 60-64.

Craik, F., & Levy, B. (1976). The concept of primary memory. In W. Estes (Ed.), *Handbook of learning and cognitive processes*. New York, NJ: Academic Press.

Gelman, R., & Gallistel, C. (1978). *The young child's understanding of number*. Cambridge, MA: Harvard University Press.

Goldmeier, E. (1972). Similarity in visually perceived forms. *Psychological Issues, 8*. (First published in 1936).

Kahneman, D., & Henik, A. (1977). Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), *Attention and Performance VI*. Hillsdale NJ: Lawrence Erlbaum.

Kimchi, R., & Palmer, S. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance, 8(4)*, 521-535.

Klahr, D. (1976). Processes for quantification. In D. Klahr & J. Wallace (Eds.), *Cognitive development, an information processing view*. Hillsdale, NJ: Lawrence Erlbaum.

Mandler, G., & Shebo, B. (1982). Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General, 111*, 1-22.

McNicol, D., & Heatcote, A. (1986). Representation of order information: An analysis of grouping effects in short-term memory. *Journal of Experimental Psychology: General, 115(1)*, 76-95.

Oeffelen, M., & Vos, P. (1982). A probabilistic model for the discrimination of visual number. *Perception and Psychophysics, 32*, 163-170.

Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.

Ryan, J. (1969). Grouping and short-term memory: different means and patterns of grouping. *Quarterly Journal of Experimental Psychology, 21*, 137-147.

Simons, D., & Langheinrich, D. (1982). What is magical about the magical number four? *Psychological Research, 44*, 283-294.

Shiffrin, R. (1976). Capacity limitations in information processing, attention and memory. In W. H. Estes (Ed.), *Handbook of learning and cognitive processes*. Vol. IV.: Attention and memory. Hillsdale NJ: Lawrence Erlbaum.

Starkey, P., & Cooper, R. (1980). Perception of numbers by human infants. *Science, 210*, 1033-1035.

Strauss, M., & Curtis, L. (1981). Infant perception of numerosity. *Child Development, 52*, 1146-1152.

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 12-31.

Wickelgren, W. (1967). Rehearsal grouping and the hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology, 19*, 97-102.

Woodworth, R., & Schlosberg, H. (1960). *Experimental Psychology*. New York, NJ: Holt.

PART III: AUTONOMY

Preface

The results described in Parts I and II point to a remarkable "smartness" of primary perception: In Part I the interpretation of the first pattern appeared to be re-used in order to reach the best representation of the sequence. In the two chapters of Part II, structural aspects of the stimulus (chapter 2) as well as a more or less metrical aspect (chapter 3) proved to be completely represented. These results were thought to support theories such as SIT that regard perception as being involved in an optimization of representation rather than a minimization of processing costs.

However, what are the limits to such optimization? Is it not always optimal to use one integral representation of as much stimuli as possible, instead of entertaining separate ones for each object or event? In other words, would not an unlimited use of context always result in the most efficient overall representation of structure? As stated before, SIT does not provide the answer. There is nothing that forbids that all codes in memory are applied in the processing of shape. On the other hand, arguments can be given why primary perception of shape would largly be context-insensitive, or *autonomous* (see the Epilogue after Part III and see chapter 4). It is clear that only empirical data can be decisive. We want to distinguish context working *backwards* (the effect of a later pattern on the interpretation of one that is presented earlier) from a more typical *forward* context. With respect to backward context, from chapter 1 we can conclude that it has only a very limited effect. After a certain critical processing time (30 milliseconds), the perceptual organization of the first pattern seemed to be fixed and could not be changed, even if a representation of the first pattern in terms of the second pattern would result in a more efficient code of the complete sequence. Although the results of chapter 2 make it questionable whether this critical time is indeed 30 milliseconds, from chapter 1 it *can* be inferred that context in the primary perception of structure can hardly work backwards. Clearly, this is a limitation to optimization. In Part III we will focus on the role of forward context: knowledge of earlier perceived patterns. With respect to the forward context effects established in the first study, one may suspect that these were specific to the experimental conditions: a very short interval between two patterns that were presented on the same spot of the screen. In contrast, the effect of context on shape perception is generally demonstrated in situations where there are clearly *two* separate and distinguishable patterns. For instance, the perceived organization of pattern B in Figure 1 of chapter 1 depends on information gathered in previous fixations on different parts of the display. This leads one to expect that such context effects are the result of prolonged processing rather than operating during primary perception. On the other hand, in many of these demonstrations there is ample time for a build-up of strong and even consciously experienced expectancies. In various recent studies efforts have been made to show that in such situations context effects are present that are genuinely perceptual (rather than mere response biases). Only a few of these studies address perceptual organization as we presently define it and not all these studies are convincing (see chapter 4). The experimental method adopted in chapters 1 and 2 of this thesis provides an excellent opportunity to investigate the effect of the context induced by a strong expectancy, while sharply discriminating effects on primary perception from biases in prolonged processing.

In the final chapter 5, an additional study on perceptual autonomy is described, although in a completely different domain, namely that of speech perception. Speech perception is also considered a "smart" process, achieving something — the recovery of linguistic information from a variable and sloppy signal — of which sophisticated computer programs are far from capable. To explain the feat of word recognition, it has been proposed that sentences are spoken within a timing structure and that this temporal context enables the perceptual system to develop adequate expectancies so that increased processing efforts are spent at the right moments. On the other hand, word-recognition may depend on the acoustical information of the word only, and the impression that sentences are rhythmically organized may be based on post-acccess processing. Again we will try to focus on early stages of processing in order to separate context effects on primary perception (the initial lexical access) from biases in later stages.

Can Perceived Shape be Primed? The Autonomy of Organization

Lucas H. M. Mens and Emanuel L. J. Leeuwenberg

University of Nijmegen, the Netherlands

To examine the role of expectancy on the organization of visual patterns, a primed tachistoscopic part-probe task was used. Sequences of a simple line figure F and a test pattern T, being a part of F, were shown, each during 10 msec with an onset asynchrony of 40 msec. Subjects had to indicate which part of F was presented separately in the sequence of F and T, by choosing among three parts shown after each trial. Subjective expectancy of alternative organizations of F was induced by showing one of three segregated versions of the F pattern (or parts of it) prior to the sequence: the priming stimulus P. As established in a pilot experiment, these versions corresponded to the preferred, to the second-best and to an unlikely organization of F, respectively. Each T pattern is a good part of one of these three organizations. The detectability of the T patterns is used to "probe" the strength of the corresponding organization of F. Subjects were aware of the fact that P and T were of the same type in 50% of the trials.

T patterns were correctly recognized more often in case of a corresponding P. However, a response-bias correction was applied, based on the distribution of the errors, discriminating a "true" effect of P and a bias due to P ánd F. P had a strong effect on the bias component, but not on the true score. Thus, no facilitation due to the prime was observed. A control experiment excluded several alternative explanations of the lack of context effect. This result suggests that early stages of shape perception are very much stimulus oriented and resistant to external knowledge. It is argued that context effects on shape perception reported in the literature are either due to a post-perceptual evaluation of stimuli or to a task simplification leading to selective attention to a local feature.

This research was supported in part by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

The authors wish to thank Martin van der Gaag for his assistance in these experiments.

We are indebted to Ar Thomassen for many helpful suggestions and incisive reading of the manuscript.

GENERAL INTRODUCTION

Common sense tells us that we can see a particular object (say, a cube) because we happen to *know* what such things (cubes) look like. This would imply a close relation between knowledge of an object and its perceived shape. In accordance, the relation between formerly acquired knowledge and shape perception has often been stressed by referring to the importance of set on the perceived organization of patterns. The flat, meaningless blobs in James's picture of a Dalmatian dog suddenly change into coherent patches of sunlight and shade when one is told that the scene is one of a dog under a tree.

Recently, efforts have been made to show that top-down effects on the perceived organization of a pattern can be genuinely perceptual, instead of merely cognitive biases. One way of doing this is to look for low-level "concomitant" effects: indirect effects of instruction or intention which are both perceptual and unexpected for the subject. For instance, when one is adapted to the vertical red grating of Figure 1a, the vertical black and white grating 1b looks green: the opponent-color aftereffect (McCollough, 1965). The

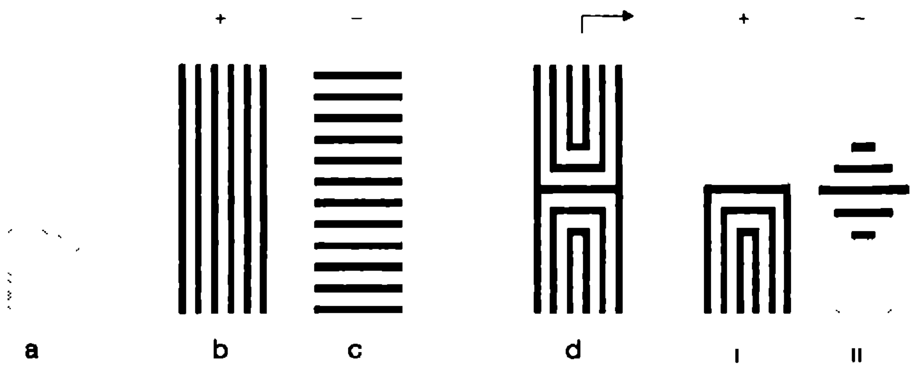


Figure 1. Adapted to red grating *a*, a black-and-white grating *b* looks green, as does *d* if one imposes organization I on it. This opponent-color aftereffect is not observed when looking at grating *c*, nor if organization II is imposed on grating *d* (Uhlarik, Pringle & Brigill, 1977). Thus, cognition seems to affect shape at a perceptual level.

aftereffect is clearly not the result of a cognitive bias. The interesting point is that the aftereffect may or may not occur depending on the pattern organization that is deliberately imposed on the black and white grating. Figure 1d supports both the U-shaped organization (vertical) as well as the rhombus organization (horizontal). The aftereffect is only observed when the U-shaped organization is dominant (Uhlarik, Pringle & Brigill, 1977). Thus, a perceptual effect coupled to pattern organization induced on a cognitive level. A similar result has been obtained by Peterson (1986) for the perceived organization of two planes in depth. These top-down effects seem to falsify what we may call the *autonomy-of-shape* hypothesis. What these studies have in common is that context effects are established for displays which are ambiguous. This is a problem. Context effects for ambiguous patterns may be considered exceptions to the overall autonomous nature of perception (Rock, 1985). More specifically, in case of ambiguity, one may assume that in an initial, autonomous stage two organizations are activated, one with and one without the concomitant effect, and that one of these two is selected afterwards in agreement with the context (Epstein & DeShazo, 1961; Hochberg, 1970, p. 102).

The goal of this study is to show an effect of context which is not explained by a post-perceptual selection, or by any other bias. We have taken several precautions in order to rule out such an explanation. One of these will be to use patterns which elicit *one clearly dominant organization*. The question then is whether context can cause an organization to become dominant that would not have been perceived without that context. An additional measure to avoid post-perceptual confounding is to use an *on-line detection task* which will be introduced in the following paragraphs.

In a seminal study, Gottschaldt (1926) introduced an embedded-figures task to investigate the effect of prior exposure on the perceptual organization of line patterns. His subjects had to report the phenomenal organization of a line pattern (the complete pattern), both with and without prior exposure to a part of the pattern, which constituted a "hidden" part of the complete pattern. See Figure 2. Prior exposure did not alter the phenomenal organization of the complete pattern: the embedded figure did not catch the eye even after many repeated exposures to it. But, more important for the present discussion, explicitly instructing the subjects to search for the parts increased the number of correct detections although again no effect of the amount of prior exposure could be established). Thus, familiarity with or frequency of exposure to a particular shape does not *completely* dictate the perceived organization (also see Djang, 1937; Hanawalt, 1942), but set *is* important.

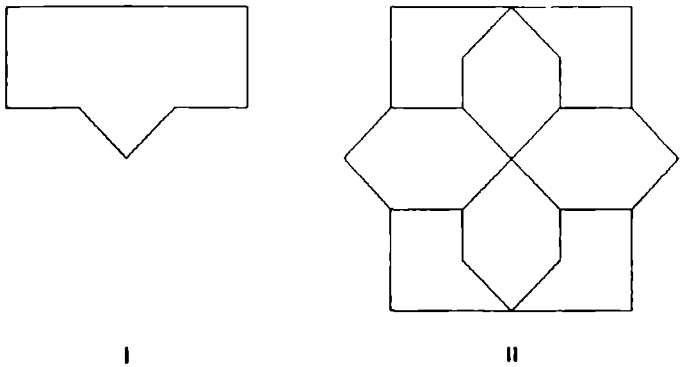


Figure 2. Depending on instruction and degree of embeddedness, a target pattern (I) may or may not be readily perceived as a part of the embedding pattern (II). Merely exposing subjects many times to pattern I does not alter the perceived organization of pattern II (Gottschaldt, 1929). However, if subjects know that pattern I may be embedded in pattern II, pattern I is detected more often, possibly because I has changed the organization of II.

However, one may raise the question to what extent these context effects are specific for the process of scrutinizing a rather complex pattern. What may be the case in the studies just mentioned is that the shape of a complete pattern (or part of it) is established in an initial autonomous phase, and that prior knowledge becomes effective only in a later phase of prolonged processing. Borrowing from the word-recognition literature (Swinney, 1970; Frauenfelder & Tyler, 1987), we intend to demonstrate context effects during the *access* phase, unconfounded by *post-access* effects. Thus, we will use simple, small patterns that can be seen "in a single glance". Furthermore, the experimental task has to be one which allows us to monitor the access phase of shape encoding, and the effect of context on it. The method to do this will be to probe the activation of specific shape organizations *on-line*. At this point we want to make a short digression to word-recognition studies where on-line tasks have received much attention.

In the *semantic priming* task, one stimulus (a word, part of a sentence) is presented shortly before a second stimulus. Facilitated processing of that second stimulus is an indication that the first stimulus has primed (lexical) entries relevant for the second one. A fine example of a study on prior knowledge and word recognition is that of Tanenhaus, Leiman and Seidenberg (1979). An ambiguous word, for instance "rose", was presented as the last word of one of two sentences: one which biased the verb reading ("They all rose") and one which biased the noun reading ("She held the rose"). After the sentence was heard, the subject had to name a visually displayed word as fast as possible. This was the test stimulus and it was compatible with the verb reading or the noun reading. Two delays between the ambiguous word and the test word were used: 0 and 200 msec. At 0 msec delay, naming latencies related to both readings were equally short, whereas at 200 msec only the reading compatible with the sentence context appeared to have survived. Thus, by measuring priming effects indirectly through the introduction of the test stimulus, it is possible to probe early in the process for specific representations (in this case word meanings), before the post-access confinements due to context are effective.

We will use the following on-line task to measure priming effects on the

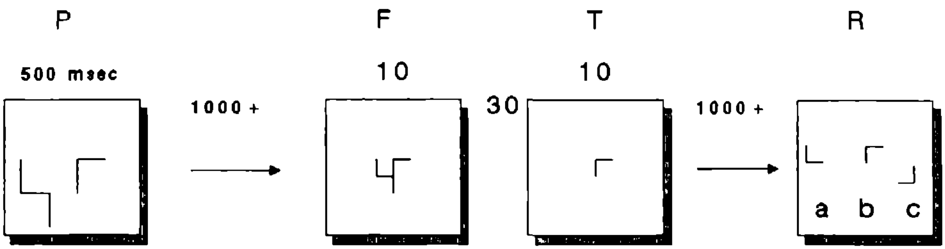


Figure 3. Example of a stimulus presentation on one trial in Experiment 2. Subjects have to indicate which one of the three T (target) patterns — *a*, *b*, or *c* — has been displayed separately in the F-T flash. The recognition of the T pattern is used as an indication what the perceived organization of F (the *focus*) is, in the context of P (the prime).

perceptual organization of a simple line pattern. Stimulus F and T will be presented in this order, each for only 10 msec with a small onset asynchrony (SOA) of 40 msec (see Figure 3). Subjects have to detect the test pattern T, a part of F. Their response will be a forced choice of one out of three alternatives (*A*, *B*, *C*). A few seconds before the tachistoscopic presentation of F, stimulus P (the *prime*) is presented. P has to be remembered and in a majority of trials it provides a valid cue as to what T will be. The idea is that the extent to which T is correctly detected may tell us whether P has imposed its organization upon F (the *focus* stimulus). F patterns are selected with the requirement that one organization is dominant (both according to empirical data as well as based on calculations of its complexity, see the Methods section of Experiment 2), and a second-best organization which is not preferred yet attractive. Accordingly, alternative types of P and T are constructed: type *A*, corresponding to the preferred organization of F, type *B*, corresponding to the second-best organization and type *C*, corresponding to an unlikely organization. A clear falsification of the autonomy-of-shape hypothesis would be found if priming of the second-best organization would result in a detection of *B*-type test patterns that is better than that of *A*-type test patterns.

We have chosen an SOA between F and T of 40 msec. In two earlier studies (Leeuwenberg, Mens & Calis, 1985; Mens & Leeuwenberg, in press) it was shown that detection of such T stimuli can indeed convey which perceptual organizations of the F pattern are currently active and that an SOA of 40 msec is about the smallest one with which effects of the shape organization can still be observed for stimuli of the type described. Thus there is evidence that with a this set-up the initial analysis of shape can be probed.

Beside ambiguity and post-perceptual confounding during prolonged processing, we wish to discuss a third hazard in studying context effects on shape organization. What appears to be an effect of context on the perceptual organization may perhaps be more properly understood as the result of *selective attention* for a specific *feature* of the test stimulus. In the present set-up, what P in fact may do, is provide a cue to the subject so that a correct response may be given on the basis of the detection of merely one local feature of T. In fact, we think this a plausible explanation of effects of prior knowledge in detection studies such as that of Bachmann and Allik (1976). (Also see Lappin & Uttal, 1976, who argue that effects of prior knowledge as such may be understood as the result of a reduction of decisional uncertainty.) It seems impossible to prime an organization of F without using P and T pairs that have simple features in common, such as line-terminators on a particular place and in a particular orientation. In other words, for practical purposes, it is impossible to separate the level of pattern-organization from the feature-level. Therefore, we will not try to preclude facilitation of T detection due to selective attention. Instead, we will start with an experiment in which we deliberately induce this effect without presenting F at all. This will provide us with a base-line against which to compare an eventual effect on the organization of F.

Finally, a fourth measure to rule out an explanation of context effects in terms of a post-perceptual process will be to submit the responses to a rigorous response-bias correction. The correction procedure is described in a separate section of Experiment 1.

EXPERIMENT 1

As has been argued in the introduction, the rationale of the present experimental set-up is that T, the third stimulus, can be used to "probe" on-line the activity of a perceptual organization of the preceding second stimulus (F). If expectancy of the prime (P), the first stimulus, imposes a particular organization on F, then we expect that detection of a T which fits into that organization is facilitated. This is, of course, a rather indirect method and some problems connected to it have been mentioned. One of these is that P may invoke *selective attention* for a specific feature of T, rather than changing the organization of F. We will first report on a preliminary Experiment which provides a base-line of selective attention effects. Such a base-line will enable us to detect and discard eventual effects of selective attention in Experiment 2.

We will introduce as many aspects of the procedure of Experiment 2 (the main experiment) as possible, without presenting the F patterns. In order to avoid a ceiling effect in the detection scores, a mask will be presented following each T stimulus. Two types of masks and two ways of presenting the P stimulus will be used (see also Figure 4). In Condition *physical match* the mask is a large number of randomly generated lines; P matches T in shape, orientation, size and location on the display. In Condition *structural match* the mask is a grid of orthogonal lines including ones that completely coincide with the lines of T (as F does in Experiment 2) plus a small number of random lines; P matches T only in shape and orientation and is larger and presented on a different part of the display, as will be made in Experiment 2. Being an exploration, no effort has been done to separate the effect of different masks and types of P patterns. Subjects will be encouraged to pay atten-

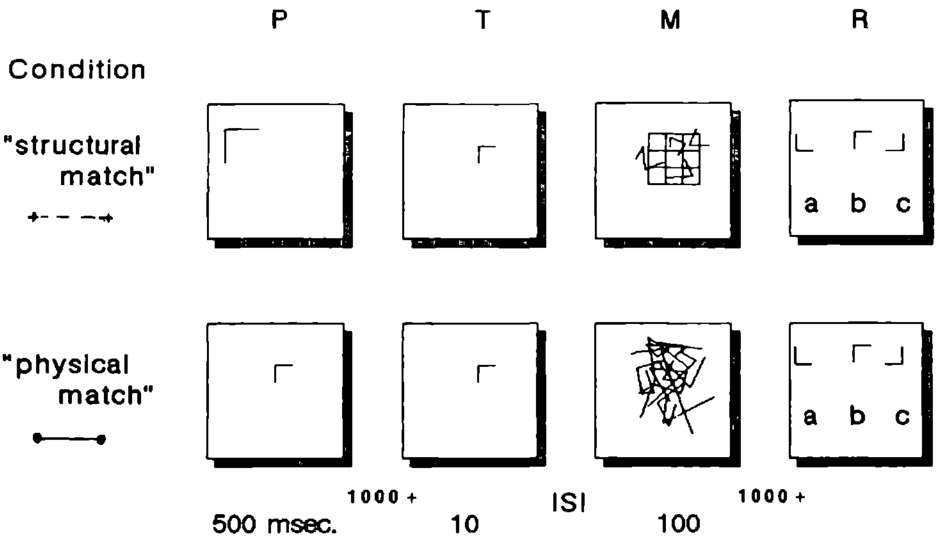


Figure 4. Figure 4. The two conditions in Experiment 1. P is identical to T in the *physical match* Condition, and larger and displaced to a side of the screen in the *structural match* Condition. Two types of M have been used as well: a grid completely aligned with T (plus some random lines), or a set of random lines (physical match). The interval between F and T is adjusted so as to achieve 30% of errors.

tion to P by telling them that P and T are of the same type in a majority of trials (actually, in 50%).

Bias correction

Of course, telling subjects that P is a valid cue as to which test pattern will be displayed will produce a considerable response bias. Furthermore, the response categories (especially in combination with the F patterns in Experiment 2) may contribute a response bias. Recently, we described a bias-correction procedure (Mens & Leeuwenberg, in press; see also the Appendix) which can be adapted to the present situation. Basically, the raw scores are separated into a probability q(i) of recognizing T(i) and a bias V(j,h) for response alternative (j) due to prime (h). Errors are assumed to result from the V-component only which makes it a high threshold model (Coombs, Dawes and Tversky, 1970). The values of the q- and V-parameters are calculated in an iterative Chi-square fitting procedure. This is done for each of the four sets (see Figure 5) of F, P and T patterns separately.

Method

Four sets of patterns were adapted from Mens and Leeuwenberg (in press); see Figure 5. Segmentations of type *A* correspond to the unambiguously preferred organization, type *B* corresponds to the non-dominant but still relatively attractive second-best organization, and type *C* to and a very unlikely organization. The reader is referred to Mens and Leeuwenberg (in press) for a description of three independent ways in which this classification has been supported. In short, these were: (a) Calculations of the *information load* of the codes of the three subdivisions follow the *A*-*B*-*C* ordering. (b) 10 Naive subjects all segmented copies of the F patterns always in the *A* fashion by preference, many times as a second choice in the *B* fashion but almost never in the *C* fashion. (c) When subjects were led to believe to be shown a sequence of an F and a T pattern, but in fact were shown an F pattern only, they showed the greatest preference for response alternative *A*, they choose *B* less often and alternative *C* still less often.

The T segments are matched with respect to the number of lines of one unit length (e.g. five lengths for the three T segments in the left-most Set in Figure 5) and differ only slightly in complexity; (see Mens and Leeuwenberg, in press, their Figure 3).

The patterns were computer generated as thin bright lines on a dark display (Vector-General Display Series 3, Model 2D3 with P4 phosphor), positioned at a distance of 2 m in front of the subject. Luminance of the patterns was adjusted to an intermediate level. Afterimages were not prominent. The visual angle subtended by a T and a P in the physical match condition was about 26 minutes of an arc. P patterns in the structural match Condition were the same pattern but twice as large. They were displaced off-centre in a random direction for maximally 4 degrees of visual angle. The room (and the display) was moderately tube-lighted. A response panel was provided with a starting key and three keys in a horizontal arrangement corresponding to three response alternatives that also appeared in a vertical array.

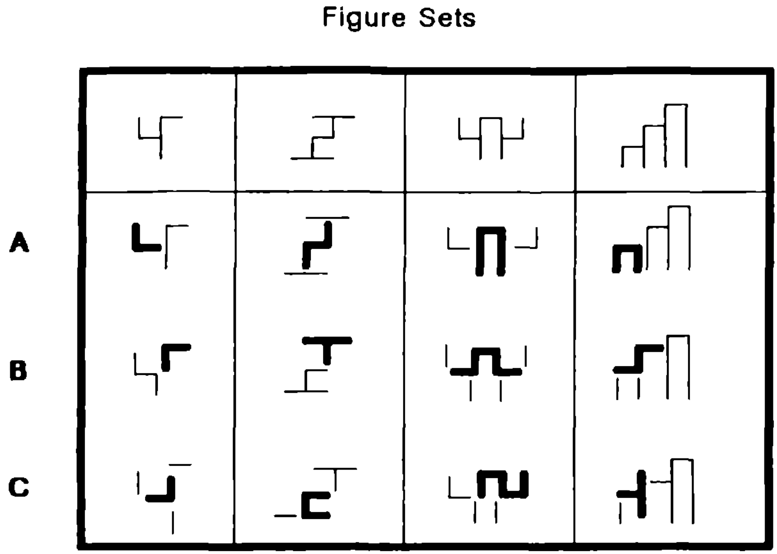


Figure 5. The four sets if line figures F (top) and three corresponding target patterns T used in the experiments. T patterns are indicated by fatter lines in this figure only. The subdivision in row *A*, *B* and *C* corresponds to the preferred, the second best and an unlikely organization of F. In Mens and Leeuwenberg (in press) the choice of the three subdivisions is motivated by referring to judgments of subjects as well as a complexity measure (Leeuwenberg, 1971).

Procedure

Thirteen students of psychology, who were paid for their participation, were shown physical match P patterns followed by random lines and ten were shown structural match P patterns and an overlapping mask. Each subject was tested individually. After a training series of 20 trials, 384 trials were administered in a random order, which took about one hour and 30 minutes. Trials were self-initiated. After the starting key was pressed, P appeared for 500 msec followed by an empty screen for 1 sec. After that second, hitting the starting key once more resulted in the presentation of T and the mask (10 and 100 msec respectively, see Figure 4). The onset asynchrony between T and the mask was initially set at 60 msec for most subjects and 90 msec for those who showed less skill during the practice session. The SOA was adapted after every ninth trial with steps of 5 msec in order to yield an error in about 1/3 of the trials. One second after the third stimulus (the mask) the three segments belonging to the set of P and T stimuli were displayed in a random order, from which the subject had to "identify the segment which was presented in the flash of two patterns, just before the grid". Except for the training series, no feedback about the correctness of the recognition of T was given. Then, at the end of each trial, Subjects had to indicate how confident they were of their response on a 5-point scale.

Each of the three T's in a set was presented 8 times in combination with each of the three P's, but T's and P's of the same type (*A* - *A*, *B* - *B*, *C* - *C*) twice as often (16 times). Thus, P's and T's were of the same type in 50% of the trials instead of 33 1/3%. In order to maximize the readiness to use P as a clue to T, subjects were told that "P is the same as T in a majority of trials and therefore can profitably be used to see T more clearly". At the end of one out of every three trials, subjects were unexpectedly confronted with a pattern which was the same as P in half of the cases and were asked to indicated whether or not this pattern was identical to P in the last trial. It was stressed in the instruction that this was a second reason to remember P.

Results and Discussion

The percentage correctly recognized T's pooled over Figure Sets and Subjects are given in Figure 6a. It is clear that T's are better recognized when preceded by a corresponding P, in both conditions. Part but not all of this facilitation can be ascribed to a bias to choose the same segment as was presented as P, as is evident in Figure 6b. The corrected recognition scores q (calculated for each Figure Set separately) are plotted in two ways. In Figure 6c the score for each P is given; in Figure 6d the scores of the two P patterns that do not correspond with a particular T are pooled. This gives a more straightforward impression of the effect of P: the corrected scores are better in case of a correspondence between P and T. This facilitation is significant both in Condition physical match ($F(1,2)=889.0$, $p<0.01$) as well as in Condition structural match ($F(1,2)=24.9$, $p<0.05$). Furthermore, it does not differ significantly between the three types of T: $F(2,6)=0.036$, $p<0.97$ in Condition physical match and $F(2,6)=1.132$, $p<0.39$ in Condition structural match.

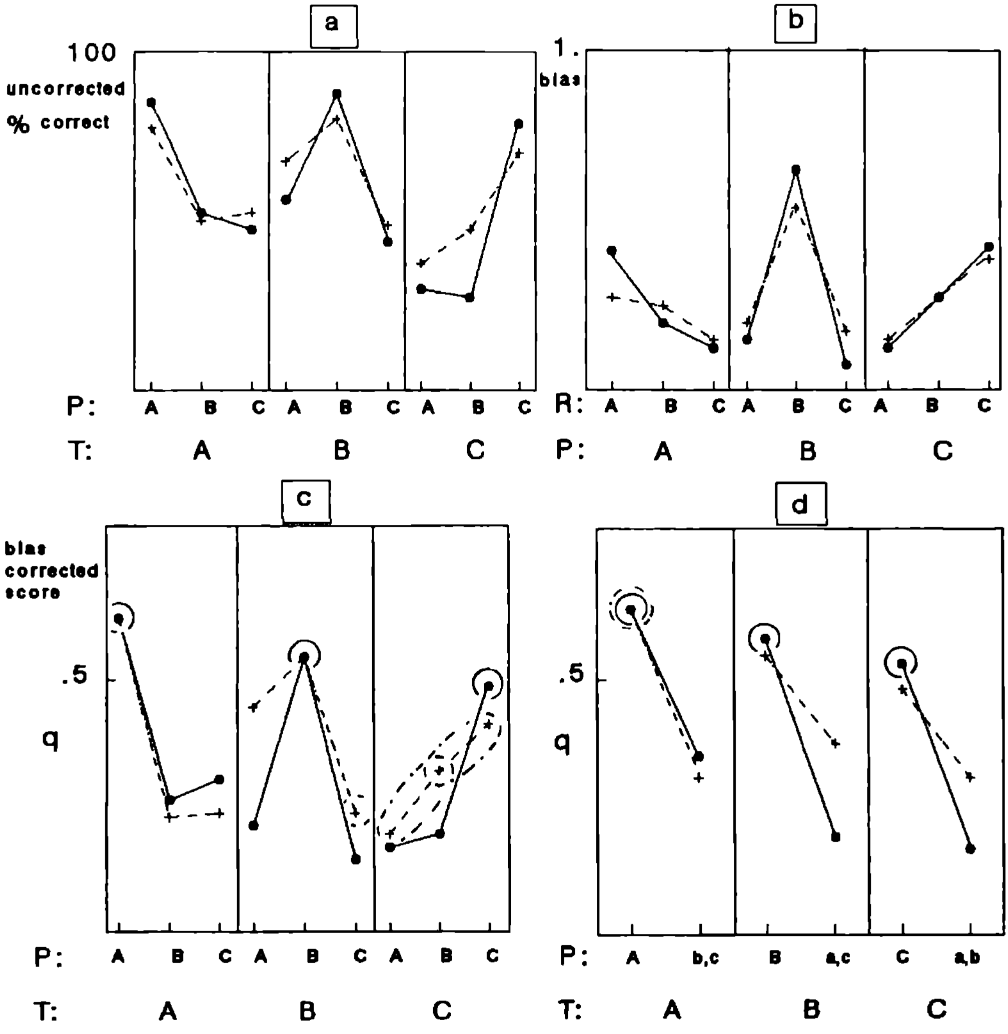


Figure 6. The results of Experiment 1, with a M stimulus, without a F pattern. 6a: Uncorrected percentage correct recognition of T. 6b: Bias values. 6c: Corrected scores, for each P separately. 6d: Corrected scores with (upper-case P) and without (lower-case P's) a correspondence between P and T. The corrected scores that are encircled differ from the other one(s) of a curve (SNK grouping, $\alpha=0.05$). As can be seen in 6d (and 6c), recognition of T is facilitated in case of a corresponding P.

A Student-Newman-Keuls multiple range test (α=0.05) has been performed on the corrected scores of each T. Scores that differ from the other (one score in 6d, two scores in 6c) are encircled. On the whole, the facilitative effect of P in the structural match Condition is clearly less pronounced.

In sum, priming facilitated the recognition of the T patterns, both when the P pattern were of the same size as T and shown at the same location on the display, and when not. It may be concluded that this task is an adequate one for inducing facilitative effects due to context.

Inspection of Figure 6c reveals that the three types of T (*A*, *B* and *C*) are not equally well detected: a Student-Newman-Keuls grouping of the *q*'s of Condition physical match averaged over P-patterns (α=0.05) shows that T type *A* is better detected than the other two. We want to explain this difference as an effect of the complexity of the T patterns. In Mens and Leeuwenberg (in press), the complexity of the T patterns is calculated using Leeuwenberg's (1971) coding model. T patterns *A*, *B* and *C* are increasingly complex according to this calculation. Although this inhomogeneity is unfortunate, we have no reason to expect that in Experiment 2 it will have an effect different from the present one. Therefore, the difference in detectability of the three sets of T patterns need not obscure possible context effects on shape perception to be found in Experiment 2.

In order to investigate whether P facilitates detection merely because it provides a cue to a simple feature of T, or whether it may also affect the perceptual organization, we will compare the facilitation obtained for the three types of T (*A*, *B*, *C*), with an F pattern (Experiment 2) and without an F pattern (Experiment 1). If there are any differences, we will relate them to the hypothesized priming effect of P on the *perceptual organization* of F.

EXPERIMENT 2

Instead of the sequence of P, T and a masking pattern of Experiment 1, we will now present P, a *focus* pattern F, and then T. The hypothesis is that the detection of T will be facilitated due to a corresponding P, not only because it has features in common with that P pattern, but also because P imposes its organization on F so that T fits into (part of) the interpretation of F and therefore is detected more easily. P, the priming stimulus, will consist of an F pattern in an "exploded" view, that is, segmented according to one of the three organizations shown in Figure 5 (with the exception of Condition 4). Subjects will be informed that in a majority of the trials, P will provide a valid clue as to which T will be contained in the sequence. Subjects will have a second reason to pay attention to P: on random occasions they will have to make a *same/different* decision about a pattern at the end of a trial which may or may not be the same as P.

In four conditions, several minor aspects of the presentation of the stimuli have been varied (see Figure 7 and the Methods Section). The most significant of these are the size and the complexity of the P pattern.

Method

Four conditions were constructed in which the relation between P and T becomes more and more direct. The P patterns in Conditions 1, 2 and 3 are derived from the F patterns by displacing target segments slightly in one direction and non-target segments in the opposite direction (see Figure 7). In the structural match Conditions (1 and 2), P was shown twice as large as F; in the physical match Conditions (3 and 4) the same size was used. In Condition 4, P was merely one of the three T patterns. In Condition 1 only, a word distractor was presented between P and the F-T sequence in order to eliminate any after-image. This was a nonsense syllable which had to be read aloud. It was not presented in the other Conditions in the light of the negative results and in order to make the task as simple as possible. A brief acoustic signal was presented in case of an incorrect target recognition in Condition 1 only.

Procedure

Nineteen students of psychology participated in Condition 1, 17 in Condition 2, 24 in Condition 3 and 7 in Condition 4; different subjects in each Condition. Stimulus presentation was basically the same as in Experiment 1, except that F-T sequences now were as follows: first F was shown for 10 msec followed by a fixed ISI of 30 msec and finally T was shown for 10 msec. The word distractor in Condition 1 was presented for 1 sec and followed P at an ISI of 100 msec. Each combination of P and T was shown 6 times, except for the combinations of P and T of the same type which were shown 12

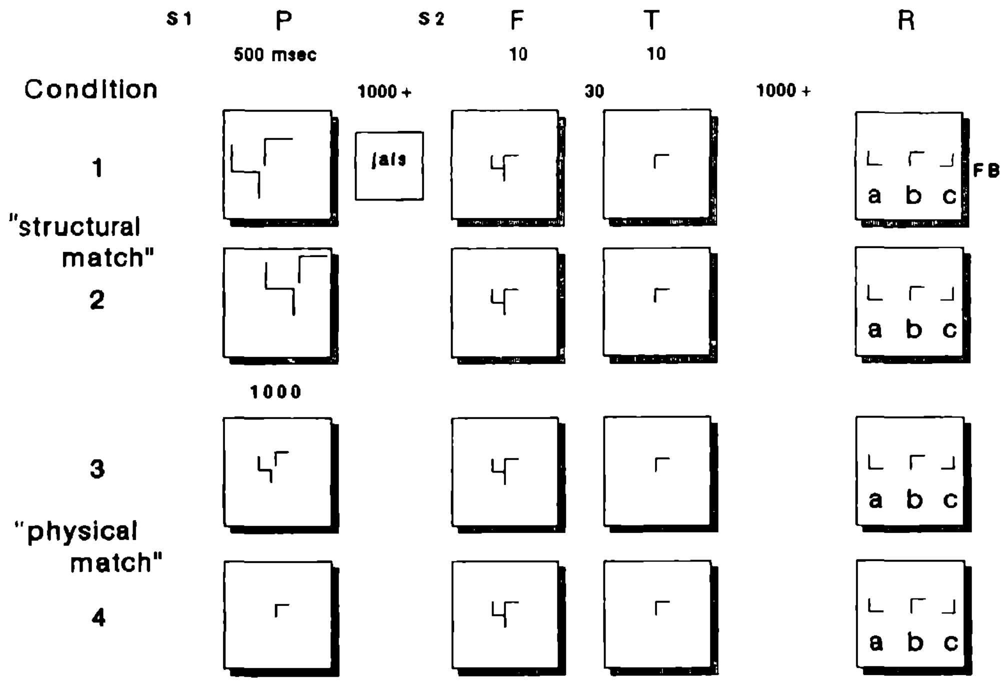


Figure 7. The four conditions in Experiment 2. Both the presentation of P and of the F-T sequence is started by the subject (S1 and S2). Feedback of errors is given in Condition 1 (FB). "Jals" is a nonsense word which has to be read aloud in Condition 1. Not displayed are a point of fixation before S1 and S2, and the occasional recognition test of P (after R).

times, in order to introduce a correlation between P and T. The resulting 288 trials were presented in a random order with the restriction that never stimuli were drawn from the same Figure Set on successive trials. As in Experiment 1, subjects were told that P could profitably be used to see the "flash" of T and the other pattern more easily. They were explicitly told that P and T corresponded in a majority of trials. Further details about the procedure are given in the Method section of Experiment 1.

Results and Discussion

In Figure 8a the percentage correctly identified target segments pooled over Figure Sets is shown for each Condition separately. Correspondence between P and T clearly yields a better recognition in all four conditions. However, these percentages correct should be submitted to the response bias correction procedure indicated in Experiment 1. The corrected "true" recognition scores q are given in 8d in a simplified way (averages of structural match Conditions interrupted lines, physical match Conditions drawn lines and also the two scores for P's that do not correspond to T averaged); in 8c for each of the four Conditions and each combination of P and T separately. The bias values are given in 8b. Student Newman Keuls (SNK, α=0.05) multiple-range tests have been performed on the corrected scores (the q values), calculated per Figure Set. No significant effect of P is obtained for the results as depicted in 8d. In Figure 8c, encircled q values differ from the other two for a particular target segment. The results of the structural match Conditions 1 and 2 have been pooled for the results as depicted in 8c because they yield highly similar (and small) effects of P. No effect of P is present in the results of the physical match Conditions.

The main result of Experiment 2 is that, contrary to Experiment 1, no significant facilitation at all due to a particular P is present in the corrected scores, as revealed by the SNK tests and shown most clearly in Figure 8d. Instead of a facilitation, the only significant effect of P, found in the combined results of Condition 1 and 2 (structural match), is a relatively low corrected score for segment *B* when P corresponds to segmentation *A*, and for segment *A* when *B* is given as priming stimulus (8c).

SNKs on corrected scores averaged over priming stimuli furthermore show a better recognition of *A* segments compared to *B* and *C* segments for the structural match results and a better recognition of *A* segments compared to *C* segments for the physical match results (again, α=0.05). A similar trend was observed in Experiment 1, though weaker.

Finally, the relation between P and the confidence with which Subjects make their response has been established for the relatively large and homogeneous set of results of Condition 3. As can be seen in Figure 8c, an increased bias does not mean less confidence (asterisks). Therefore, subjects did not display a deliberate strategy to choose the P-corresponding segment if they were uncertain.

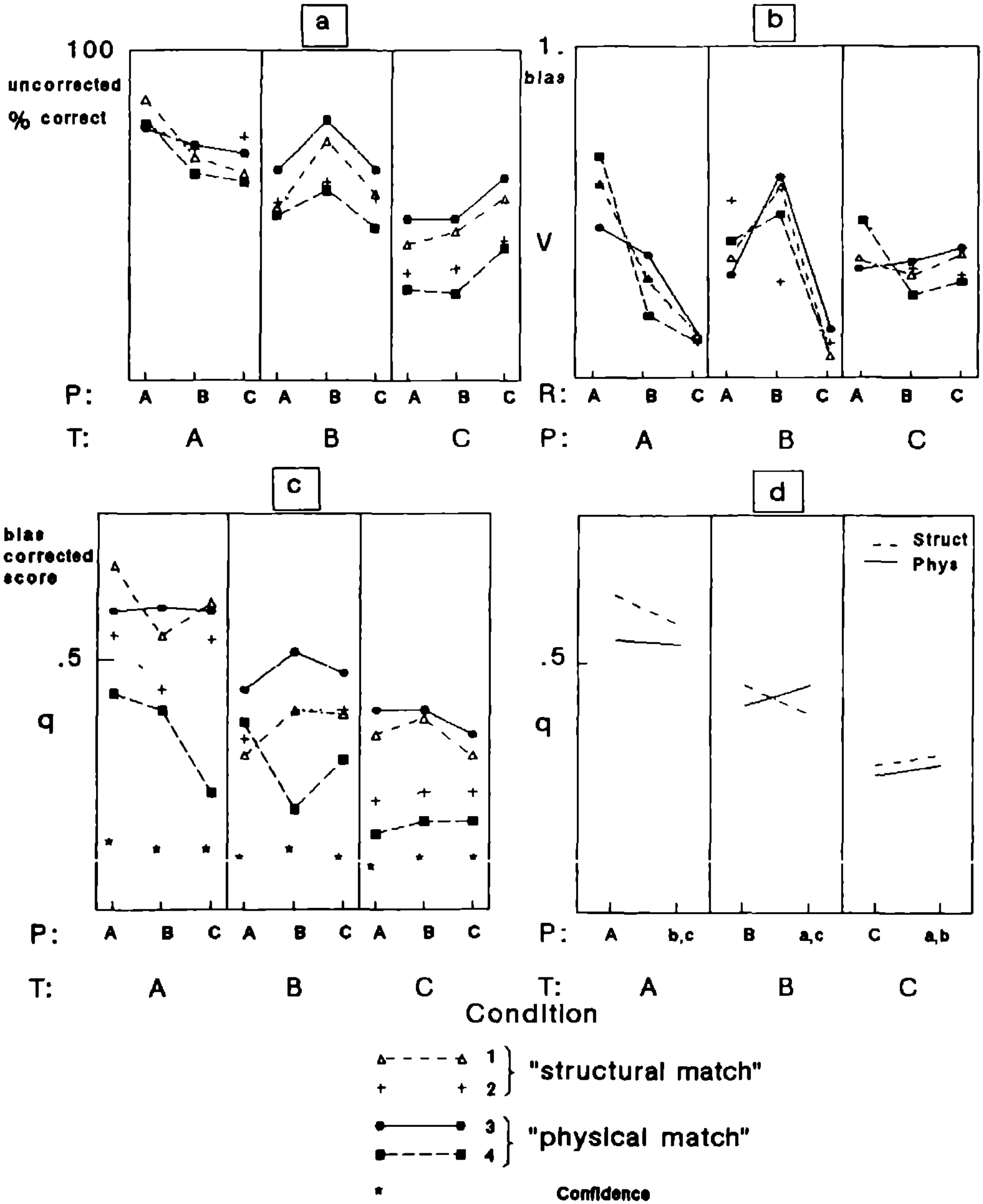


Figure 8. The results of Experiment 2, with an F stimulus. In contrast to Experiment 1, no facilitation due to a correspondence between P and T is observed (8d and 8c). Instead, in 8c inhibitory effects of P seem present.

GENERAL DISCUSSION

An important point for the interpretation of the results is whether we have succeeded in creating an effective context, regardless what that context has meant for the organization of the F pattern. This seems to be the case. Our subjects did expect to be shown a test pattern (T) that was the same as the prime (P). The expectation became apparent in three ways: (a) In Experiment 1 *selective attention*, cued by the P's, facilitated the detection of the T patterns. (b) In both Experiment 1 and 2 a strong response bias appeared in favor of T patterns of the same type as P. (c) Not only were subjects prepared for a particular T on the basis of P, they also appeared to *see* such a T, even if it was not shown: T patterns of the same type as the P were chosen with an increased confidence. Yet, the main finding is that after correction for response biases, no indications were found that the P patterns primed a particular perceptual organization of the *focus* patterns, F (Experiment 2). What definitely did not happen was that the non-dominant (but relatively good) second-best organization, *B*, became dominant if primed.

Before we discuss the fact that P did not facilitate the detection of corresponding T's at all in Experiment 2, two issues have to be dealt with.

The first issue is the apparent *inhibition* observed in the structural match conditions of Experiment 2: P's priming the preferred organization, *A*, inhibited the detection of the second-best T patterns, and, vice versa, P's of type *B* inhibited T patterns of type *A* (encircled points in Figure 8c). The inhibition did not appear in the physical match conditions, which might indicate that physical match primes were less effective in biasing a particular organization. Still, the inhibition could be an unstable and unimportant effect and could perhaps even be an artifact. If it is not to be considered an artefact, the only sensible interpretation that we can make leads to the conclusion that it is a *post-access* effect. The reasoning runs as follows: In an initial access-phase, organizations *A* ánd *B* are generated of the F pattern, but not organization *C* (see Mens and Leeuwenberg, in press). In a later phase (post-access), a check is made of the consistency of the prime and the perceptual interpretation (similar notions for word recognition are Neely's "associative-matching" strategy, 1976, and De Groot's "coherence check", 1983). Apparently, a match between the prime and *one* of the perceptual alternatives leads to the active suppression of the *other* alternative, while a complete mismatch has no consequence at all. Another argument to conclude that the inhibition is a post-access effect is that a genuinely perceptual effect of priming (e.g. via spreading activation) can be expected to be primarily facilitative. However, the present effect of the P patterns is not facilitative, it is "costs-only". As such, it is more natural to assign it to a decisional component. So, instead of indicating a context effect, the inhibition can be understood as an extra indication of autonomy.

The second issue is the following. Did we adequately measure the perceptual organization of F? In other words: Is the detection of the T segments affected by the organization of F? We want to give two arguments that this was the case. (a) Using the same patterns (but without a P), task and ISI, Mens and Leeuwenberg (in press) were able to infer from the detection of T patterns a specific property of the shape interpretation of F, namely that the F patterns give rise not only to the preferred, but also to the second-best organization. Specifically, in that study *A* and *B* organizations of F appeared to be *present*, although *A* was more strongly *preferred*. (b) The interpretation of the inhibition effect given above depends on the assumption that P patterns *A* and *B* related to two different organizations *of the F pattern*. If the perceived organization(s) of F are not important, similar results as in Experiment 1 should be obtained.

Now we will consider three more or less methodological objections that one may raise against the present conclusion.

The bias correction may have been excessively effective, eliminating "true" effects of P. The bias correction procedure is based on a high threshold model which means that errors are assumed to be made on the basis of the bias component alone (Coombs, Dawes & Tversky, 1970). This assumption may not have been met. However, the notion of overcorrection does not seem to apply. The data of Experiment 1 have been subjected to exactly the same analysis as those of Experiment 2 and this did not eliminate what we interpreted as the effect of *selective attention*. Moreover, another effect was visible in the results of Experiment 2, even after correction, namely the overall superiority of detection of *A* over *B*, and of *B* over *C* test patterns. In an earlier study (Mens & Leeuwenberg, in press), several arguments for the validity of the same correction procedure, albeit in a simpler form (not accounting for a bias due to primes), are put forward.

The abrupt and ultra-short presentation may have induced many effects not relevant to normal perception (Haber, 1983). For instance, it may have disrupted the normal influx of prior information. Causes of such a disruption might be a transient-on-sustained inhibition (Breitmeyer & Ganz, 1976) or a startle response (Hochberg, 1970). However, an explanation in terms of disruption is hardly compatible with the fact that P in Experiment 1 did have a facilitating effect. Moreover, a tachistoscopic presentation seems to be essential and unavoidable if one wants to study early processing phases, as we have argued in the introduction.

Is the lack of context effect specific for meaningless patterns? Since prior knowledge from *within* the shape domain does not seem to have an effect, it seems unlikely that knowledge from a higher level will. Still, one might wonder whether the anticipation of an object that has a "real" meaning for the perceiver cán be effective. There are two sets of data which show that such an anticipation in general seems to result only in a more enriched (and perhaps speeded) "semantic" (higher-level) classification instead of in a *change* at the level of shape organization.

(a) Riddoch and Humphreys (1987) described a case of Optic Agnosia. Optic Agnosia underlines the functional disjunction of knowledge about shape or "object-structure" and higher-level knowledge or "object-meaning". Patients displaying Optic Agnosia ar capable of making adequate gestural responses to objects, indicating that their perception of the shape of these objects is unimpaired, but they cannot *name* these perceived objects, nor give any functional association to them. At the same time, object knowledge in general is intact, evidenced by the fact that it cán be accessed via other modalities (e.g. language).

(b) Normally, shape does not change at all, even if knowledge about object-meaning keeps accumulating. In other words, an initial "semantics-free" shape representation generally remains unchanged. Witkin and Tenenbaum (1983), for example, argue that a picture of the rings of Saturnus does not look different after learning what these rings are made of: "(..) What is remarkable is the degree to which such naively perceived structure survives more or less intact once a semantic context is established (..)" (p. 481). Similarly, when we know that we experience a "visual illusion", the initial shape conflicts with what other sources of information tell us that we *should* see, but again the initial shape survives. This suggests that normally, without a conflict, the initial organization is incorporated in, rather than replaced by later interpretations. Many remarkable illusions are construed by Kanisza (1985), who poses that "the visual system resolves its 'problems' without regard to logic, expectations, and knowledge" (p. 31).

REFERENCES

Bachmann, T., & Allik, J. (1976). Integration and interruption in the masking of form by form. *Perception, 5*, 79-97.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review, 83*, 1-36.

Coombs, C., Dawes, R., & Turvey, A. (1970). *Mathematical Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

De Groot, A. M. B. (1983). *Lexical-context effects in visual word recognition*. Unpublished doctoral dissertation, University of Nijmegen.

Djang, S.-S. (1937). *Journal of Experimental Psychology, 20*, 29-59.

Epstein, W., & DeShazo, D. (1961). Recency as a function of perceptual oscillation. *American Journal of Psychology, 74*, 215-223.

Frauenfelder, U. H., Komisarjevsky Tyler, L. (1987). The process of spoken word recognition: An introduction. *Cognition, 25*, 1-20.

Gottschaldt, K. (1926). Gestalt factors and repetition. In W. D. Ellis (Ed.), *A source book of Gestalt psychology*. pp. 109-135.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *The Behavioral and Brain Sciences, 6*, 1-54.

Hanawalt, N. G. (1942). The effect of practice upon the perception of simple designs masked by more complex figures. *Journal of Experimental Psychology, 31*, 134-148.

Hochberg, J. (1970). Attention, organization,and consciousness. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts.

Kanisza, G. (1985). Seeing and knowing. *Acta Psychologica, 59*, 3-23.

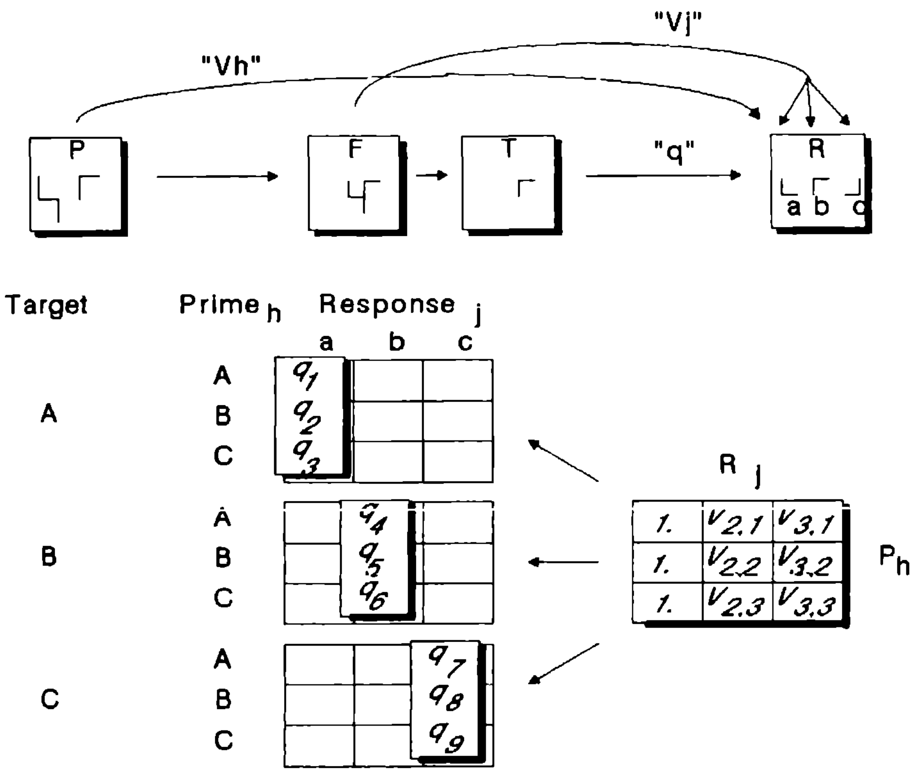
Lappin, J. S., & Uttal, W. R. (1976). Does prior knowledge facilitate the detection of visual targets in random noise? *Perception and Psychophysics, 20*, 367-374.

Leeuwenberg, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology, 84*, 307-349.

Leeuwenberg, E., Mens, L., & Calis, G. (1985). Knowledge within perception: masking caused by incompatible interpretations. *Acta Psychologica*, *84*, 91-103.
McCollough, C. (1965). Color adaptation of edge-detectors in the human visual system. *Science*, *149*, 1115-1116.
Mens, L., & Leeuwenberg, E. (in press). Hidden figures are ever present. *Journal of Experimental Psychology: Human Perception and Performance.*
Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory and Cognition*, *4*, 648-654.
Peterson, M. A. (1986). Illusory concomitant motion in ambiguous stereograms: Evidence for nonstimulus contributions to perceptual organization. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 50-60.
Riddoch, M. J., Humphreys, G. W. (1987). Visual object processing in optic aphasia: A case of semantic access agnosia. *Cognitive Neuropsychology*, *4*, 131-185.
Rock, I. (1985). Perception and Knowledge. *Acta Psychologica*, *59*, 3-22.
Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behaviour*, *14*, 645-660.
Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, *18*, 427-440.
Uhlarik, J., Pringle, R., & Brigill, M. (1977). Color aftereffects contingent on perceptual organization. *Perception & Psychophysics*, *22*, 506-510.
Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. I. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision*. New York, NJ: Academic Press.

APPENDIX

The percent correct recognition of T is separated in two components: the "true" detection (q) and the bias V(j,h) due to P(j) and response category R(h). In an iterative Chi-square minimization (program "Minuit", Cern, Geneva) best fitting q and V values are found: one q for each T pattern of a Figure Set in the context of nine combinations of P and R, and one V for each combination of P and R. Bias values are relative, so the bias for R(1) has been given a fixed value.



$$\text{Correct} = \left\{ q_i + (1-q_i) * V_{j,h} / V_{\text{sum}_h} \right\}$$

$$\text{Error} = \left\{ (1-q_i) * V_{j,h} / V_{\text{sum}_h} \right\}$$

Evidence against a Predictive Role for Rhythm in Speech Perception

Lucas H. Mens and Dirk-Jan Povel

University of Nijmegen, The Netherlands

Some languages create the impression of being stress timed Claims have been made that this timing of stressed syllables enables the listener to predict the future locations of informative parts later in a sentence The fact that phoneme monitoring is delayed when targets in a spoken sentence are displaced has been taken as supporting this claim (Meltzer, Martin, Bergfeld Mills, Imhoff and Zohar, 1976). In the present study temporal displacement was induced without introducing phonetic discontinuities. In Dutch sentences a word just in advance of a target-bearing word was replaced by another one differing in length. Results show that the temporal displacement per se did not have any effect on phoneme-monitoring reaction times. Implications for a theory of speech processing are discussed.

GENERAL INTRODUCTION

In many instances speech perception depends on the perceived relative timing of acoustic events. Thus it is well established that at the local level the relative onsets of adjacent events may determine the phonemic identity of certain speech segments. The role and even the existence of global timing, however, is controversial. Nevertheless, the impression that English and several other languages are so-called "stress timed" pervades a considerable number of studies of speech production and perception. Stress timing implies some sort of temporal structuring (rhythm) of the syllables carrying major stress in a sentence.

Requests for reprints should be sent to L H Mens, P O Box 9104, 6500 HE Nijmegen, The Netherlands.

This research was supported in part by the Netherlands Organization for the Advancement of Pure Research (Z W O).

The authors wish to thank William Marslen-Wilson, Christopher Darwin, Anne Cutler, an anonymous reviewer and especially Steve Handel for their constructive criticism and Hans Jansen and Chris Bouwhuisen for their technical assistance

The most straightforward notion of stress timing is that of isochrony or equidistant timing of stressed speech events. It has been difficult to test this notion by physical measurements in part due to the uncertain localization of the center of the stressed syllable (Haggard, 1971; Lehiste, 1973; Ohala, 1975). For this reason, the possible role of isochrony in *perception* has been investigated (Allen, 1972; Donovan and Darwin, 1979). The results are inconclusive. As an example, the reproductions of perceived accent structures tend toward isochrony both for English (Donovan and Darwin, 1979), a stress-timed language, and for French, which is not regarded as a stress-timed language.[1] It may be noted that isochrony seems not to be beyond the control of the speaker since when specifically asked to, speakers can control the timing of syllables with major stress (Fowler, 1983), and this controlled timing can be of perceptual relevance (Huggins, 1975; Fowler, 1981, 1983).

In line with other studies, here rhythm in speech is taken to imply predictability. As in music, rhythm would impose well-defined and predictable timing constraints enabling the anticipation of future events of speech. Because of this predictability, not only a cue to the status of those elements (accented—not accented) but also a means to direct processing resources would be provided. Martin (1972) and Meltzer et al. (1976) developed a theory of perception that "describe[s] the listener in part as a sort of dynamic feed forward device, performing analysis by synthesis in a real time, anticipatory fashion, and who, given cues from early elements in an utterance, can (in effect) predict future vocal-tract movements of the speaker" (1976, p. 277).

To test these ideas, Meltzer et al. (1976) assessed the effect of distortion of the temporal structure on sentence processing by temporally manipulating the signal in front of the target. Phoneme reaction times in three main conditions were compared to reaction times to the original recordings: "early", "late intact" and "on time". The three manipulated versions are shown in Table I using parts of one of Meltzer et al.'s sentences: "He laughed and laughed till his belly wiggled like jelly". To create the "early" version, 100 msec of the material immediately in front of the target (phoneme /b/) was spliced out. To create the "late intact" version 100 msec of noise was inserted immediately in front of the target. The "on time" version maintained the same timing as the normal version, but 100 msec of white noise replaced the original signal in front of the target. The results offered no straightforward interpretation: reaction times for the early and on-time versions were

[1] S Isard, D Scott and B. de Boysson-Bardies. The perceived rhythm of English and French as assessed by the tapping task Paper presented at the 10th meeting of ASA, Norfolk, Va , May, 1984.

Table I

Sentences Used by Meltzer et al. (1976) and Results

Condition	Sentence	Reaction time[a]
normal:	... till his Belly ...	620
early:	... till hiBelly ...	700
late intact:	... till his Belly ...	562
on time:	... till hi*Belly ...	630

[a] In msec

Note A representative selection of the 17 conditions in the original paper is shown The reaction time for the "on time" condition was estimated on the basis of comparable "early" conditions in Experiments 1 and 2 and Experiment 4 of the same paper White noise is marked with an asterisk

longer than for the normal version, but the late-intact version did not create a decrement. However, Meltzer et al. interpreted these outcomes as the result of deceiving the rhythm-based anticipation of accents. Although he considered that any aspect of the signal that departs from what a listener expects to be normal in an utterance most probably will have a disruptive effect on sentence processing, Meltzer et al. stressed that some of the late-intact conditions resulted in faster reaction times than did the normal sentences, "and hence [that] intervention effects by themselves probably cannot provide a sufficient account for slower RT" (1976, p. 278). Anyway, all that can be said is that apparently in these cases facilitating effects were greater than disruptive effects. The important point to note is that in all modified sentences used by Meltzer et al., the manipulation of the experimental variable (temporal displacement) was confounded with a distortion of the phonetic material.

The goal in this study was to test the anticipation hypothesis without introducing segmental distortions. In Experiment 1 the effect of pure temporal disruption on phoneme monitoring RT was examined. Buxton (1983) conducted an experiment very similar to the present one and found indications of rhythm, as did Meltzer et al. An exposition of her study will be postponed to the General Discussion. In Experiment 2 temporal displacement was compared to distortions of the acoustic signal in order to make the results more comparable to those of Meltzer et al.

EXPERIMENT 1

If the temporal structure of the accents in a spoken sentence facilitates processing of the accented items, then an artificial displacement of an accent later in the sentence should become apparent as a relative setback generating slower phoneme reaction times. In this experiment, the most important temporal modifications were brought about by replacing the pretarget word with another one of different length, taking care that this manipulation did not distort any other feature.

Method

Subjects

Eighty-one students of psychology, about half of them female, half of them male, participated in the Experiment to fulfil their introductory course requirements

Sentence Material

Twenty-four pairs of sentences were spoken by a female speaker and recorded on audiotape. Each pair consisted of two identical sentences except for the pretarget word, which was short in one version and long in the other All short pretarget words contain one strong syllable, 19 of the 24 long pretarget words contain one weak and one strong syllable, 2 contain two weak and one strong syllable, whereas only 3 contain two strong syllables. In composing the sentences, care was taken to maintain approximately the same level of semantic continuity in both versions. The speaker was instructed to keep the intonation of both versions as equal as possible and to maintain the same overall speaking rate. The target phoneme /b/ occurred only in word-initial position In 19 of the 24 cases the target word was definitely accent-bearing.

Precautions were taken to avoid any regularities that might signal the approach of the target First, the pairs of pretarget words occurred with equal frequency in a representative sample of speech (Uit den Boogaart, 1975). Second, while restricting ourselves to words beginning and ending with one of the stop-consonants /p/, /t/ or /k/ for easy splicing, the 24 pairs of pretarget words as a whole mirrored the frequency distribution of words beginning with /p/, /t/ or /k/ as well as the distribution of nouns and verbs. Word-final stop-consonants of the verbs were all /t/, but this being the most frequent case for present-tense third-person verbs in Dutch, it should not provide a significant cue to the listener. Third, the predictability of the target was kept low by using rather long sentences. Meltzer et al. (1976) located their targets 0 to 8 syllables from the end of the sentences, which were about 18 syllables in length. In our case, location of the target varied from 3 to 16 syllables from the end in sentences with an average length of 22 syllables.

The six conditions used are shown in Table II. The two original recordings made up Conditions 1 and 2. Starting with these two recordings, four other versions were constructed using digital signal processing techniques. Cutting points for word splicing were chosen just before the stop-consonant burst at the beginning of the pretarget word, and after this word just before the vocal murmur of the target. Two versions were formed by exchanging the pretarget word Condition 3 inherited its sentence frame from Condition 1 and its pretarget word from Condition 2. Condition 4 was formed in a likewise manner from Conditions 2 and 1 Conditions 5 and 6 both were derived from version 4. In Condition 5 just enough silence was spliced in between the short pretarget word to let the target word start at the same moment in time as it did in the original long pretarget versions (Condition 2). Finally, in Condition 6 the /b/

Table II

Sentences, Manipulations and Results in Experiment 1

Condition	Sentence	RT[a]	S.E.
1 normal short	. . stukje kurk bovenop de . . .	604.5	9.89
2 normal long	. . STUKJE KNOFLOOK BOVENOP DE . . .	566.4	9.82
3 displaced long	. . stukje KNOFLOOK bovenop de . . .	574.5	9.20
4 displaced short	. . STUKJE kurk BOVENOP DE . . .	604.4	10.80
5 silence on time	. . STUKJE kurk BOVENOP DE . . .	579.4	10.72
6 silence later	. . STUKJE kurk BOVENOP DE . . .	587.5	10.93

[a] In msec (averages of about 260 observations).

Note Abbreviated example of one sentence, showing the relevant fragment and its subsequent manipulations Target phoneme is /b/. Upper and lower case are used to depict the two different recordings. Sentences are in Dutch (see Appendix) was preceded by two times the silence duration. These conditions were added in order to relate to Meltzer et al.'s condition "late-intact".

The manipulated sentences sounded perfectly normal both with respect to articulation and intonation. A naive listener who was asked to indicate any irregularities listened to versions 1 to 4 of all 24 sentences. He reported six cases of sloppy pronunciations of consonants, all occurring at locations where no splicing had been performed, while in only one case splicing seemed to have resulted in an abnormally long pause before a /b/ target (sentence 11, version 4).

Procedure

All 144 experimental sentences plus 24 filler sentences (similar to the experimental sentences in all respects except lacking a /b/) were stored on a RKO7 disk after digitization (sample frequency 10 kHz, 24 dB/oct at 4,000 Hz), which provided random access. Stimulus presentation and response registration were controlled by means of a PDP 11/45 computer. The 144 (6 conditions times 24) sentences were split up into six blocks, each block containing only one of the six versions of a sentence. Each subject, run individually, was presented one block plus filler sentences in a unique random order. A set of 18 specially recorded sentences (half of them target-bearing) was administered for training purposes.

The subject started each trial by pushing a key. The main task was to react as quickly as possible upon hearing a /b/ by pushing another key. To make the task linguistically more valid, the subject was also required to paraphrase the sentence. A microphone was placed in front of the subject to give the impression of being recorded.

Response latencies were measured in msec from the beginning of the voice-onset time of the target phoneme /b/. This location was established prior to the experiment with an accuracy of 10 msec. Sentences yielding reaction times beyond 1,100 msec and misses were repeated after all remaining ones. The experimenter immediately warned the subject in case of a miss by naming the target-bearing word. Fifteen of the initial 81 subjects were excluded because they missed the target more than 10 times. Thus 66 subjects remained, 11 subjects in each of the 6 groups, each group assigned to one of the 6 different blocks of sentences.

Results

Introducing a temporal distortion by exchanging long and short pretarget words had no effect: Conditions 1 and 4 and Conditions 2 and 3 yield very similar reaction times (see Table II). A two-way analysis of variance on the raw data of Conditions 1 to 6 showed significant main effects of Conditions [$F(5, 1{,}397) = 2.652$, $p < 0.05$] and Sentences [$F(23, 1{,}397) = 5.484$, $p < 0.001$). The interaction was significant [$F(115, 1{,}397) = 1.255$, $p < 0.05$] but can be ascribed to Conditions 5 and 6: Analysis of Conditions 1 to 4 alone yielded main effects of Conditions [$F(3, 95) = 4.589$, $p < 0.01$] and Sentences [$F(23, 95) = 4.865$, $p < 0.001$] but no interaction [$F(69, 95) = 1.202$, $p > 0.1$]. The lack of effect is supported by a multiple range test (lsd procedure of SPSS, $\alpha = 0.05$): RTs in Conditions 1 and 4 on the one hand and 2 and 3 on the other do not differ. Though irrelevant to rhythm, an effect of the length of the pretarget word is clearly present (same lsd procedure, $p < 0.05$ for the differences between 1–2, 1–3 and 3–4): Long pretarget words elicited shorter RTs. Conditions 5 and 6 did not differ from each other or from Conditions 1 to 4.

Additional Analyses

First, the sentences were divided into one set containing the target-bearing words that are clearly accented ($n = 16$) and a second set ($n = 8$) containing the less clearly accented words. A t-test revealed that the average RTs ($m1 = 567.4$; $m2 = 624.4$) differed significantly ($p < 0.01$). Second, a division with as criterion long versus short target-bearing words did not yield an effect ($m1 = 594.2$ and $m2 = 580.8$, $t = 0.719$). Third, no relation was found between frequency of the pretarget word and reaction times (Conditions 5 and 6 were excluded from this analysis as they are contaminated because of clear phonetic disruption) [$F = 1.613$, $p > 0.1$]. Fourth, the number of times the target in a sentence was missed correlated significantly with the mean latency of correct responses to the target ($r = 0.69$), showing that difficult sentences were reacted to more slowly and were often missed. The number of misses was spread evenly over the 6 conditions.

Discussion

The main result of this experiment is that a deformation of the temporal structure of a sentence without concurrent phonetic distortions does *not* lead to slower detection of a target phoneme just after the deformation. This result calls into question the contention that a perceived accent structure allows for efficient look-ahead processing of forthcoming elements.

In a survey by Cutler and Norris (1979), several effects that typically occur in phoneme-monitoring experiments are listed. Two of these effects are found in the present results. First, we found that Conditions 2 and 3 with long pretarget words produced significantly faster responses than did short pretarget words (Conditions 1 and 4). Second, stressed target words evoked faster reaction times than moderately stressed words. A third effect mentioned by Cutler and Norris could not be established: we did not find the frequently reported faster RTs on short target-bearing words. This can easily be understood because of the disproportionately large number of function words with small semantic focus and perhaps a lower level of stress in the group of short target-bearing words.

EXPERIMENT 2

Temporal and Phonetic Alterations

In the previous experiment two conditions examined by Meltzer et al. (1976) were not included. These were the conditions named "on time" (noise added) and "early" (material deleted). In Experiment 2 we replicated these conditions. According to our interpretation of the effects, the noise should cause a negative effect due to the disruption of processing but a positive effect due to providing an anticipation combining to yield a small or zero effect. The early condition, in contrast, is expected to result in a retardation of reaction times due to the deletion of pretarget information (and not due to altering the perceived accent structure, as argued by Meltzer et al.). If replicated, these results would give extra evidence that Meltzer et al.'s and our findings are comparable and that differences in the outcome cannot be attributed to gross differences in materials.

Method

Subjects

Thirty students from the same population as in the first experiment were tested in groups of about four.

Sentence Material

This experiment comprised three conditions. The first version of the sentences of the previous experiment (Condition 1, short pretarget word) formed Condition 1 ("normal"). Conditions 2 ("early") and 3 ("noise") were constructed starting with version 1 and employing the same techniques as discussed previously. In Condition 2 a 100-msec speech fragment just in front of the target was spliced out, and the two remaining parts were abutted, resulting in a 100-msec advancement of the target. In Condition 3 this same 100-msec portion was not deleted but was replaced by 100-msec white noise (see Table III). The amplitude of the noise roughly matched the loudness of a stressed vowel. To prevent the noise from becoming a reliable target cue, in six of the 24 filler sentences noise was inserted at locations comparable to those in the experimental sentences.

The sentences were recorded on tape. Three experimental tapes were constructed, each tape containing on one track a randomization of all 24 sentences (each in only one of its three versions) and the filler sentences (so instead of the six blocks of the previous experiment, now three blocks were composed) A silence of about five seconds separated trials.

Procedure

The task was the same as in the former experiment, namely to react by pushing a button as quickly as possible on hearing a /b/. In the secondary task, subjects

Table III

Sentences, Manipulations and Results in Experiment 2

Condition	Sentence	RT[a]	S.E.	N misses
1	normal . . kurk bovenop . .	541.0	9.37	10
2	early . . kurbovenop . .	595.4	10.90	40
3	noise . . kur**bovenop . .	545.7	9.94	21

[a] In msec (averages of about 200 observations in each condition).

Note: In "early", the target is 100 msec advanced; in "noise", 100 msec of white noise (asterisks) replaces pretarget material.

were now told that they would be presented with a written recognition test after the experiment to test their understanding. Prior to the experiment, a practice tape of 16 trials (half of them with a target) was administered, followed by a recognition test.

Presentation of the sentences was by headphones. Subjects sat separately with a response panel and a video display in front. In case of reaction times faster than 100 or slower than 1,000 msec, "error!" was immediately displayed on the video. Response collection was controlled by means of a PDP 11/34 computer.

Results

Overall, Conditions had a strong effect on RT [$F(2,646)=8.672$, $p<0.001$] as shown in Table III. This effect is due to a significant retardation of reaction times in the condition "early" as compared to the "normal" and "noise" conditions. "Normal" and "noise" did not differ, LSD procedure, $\alpha=0.05$.

Other Analyses

The same division of the 24 sentences into two sets according to accent level, as in the previous experiment, again yielded significant results ($p<0.05$). Sentences that often failed to produce a reaction were reacted to more slowly ($p<0.01$) by the subjects who did respond.

Discussion

The results confirmed the prediction that advancing the target by deletion of pretarget material slowed RT. In fact, RTs to "early" sentences were 55 msec slower than RTs to normal sentences, close to the 60-msec difference that Martin reported. Moreover, there was a high number of misses (40), also implying that processing was disrupted.

The effect of noise just before a target appeared to be inconsistent. Although the number of misses (21) was twice as large as for normal sentences, mean and variability of RTs was hardly increased. This suggests that the search processes responsible for /b/ detection do not have direct access to the "raw" acoustic input. Instead, a perceptual separation of speech and noise seems an obligatory first stage. If it succeeds, normal detection is possible; if it fails, total ineffectiveness of the (distorted) /b/ features is the result. The separation appeared to be quite powerful: several subjects showed great surprise at the end of a session when asked about the 14 sentences with a noise burst. Although told to expect these, they had been totally unaware of extraneous sounds. This noninterference of extraneous sounds with perception of speech is in line with the "phonemic restoration" effect reported by Warren and Obusek (1971).

GENERAL DISCUSSION

The results presented here clearly indicate that temporal distortions by themselves have *no* effect on phoneme-monitoring reaction-times in a spoken sentence. The effects reported by Meltzer et al. therefore most probably resulted from phonetic distortions such as noise substitution, silence insertion and signal elimination in their speech material.

On this basis, we would suggest that spoken sentences do not possess global rhythmical structures guiding the processing of the signal. Rather, we would contend that the production of speech is basically an incremental concatenated process in which the durations of speech events are determined by temporal constraints of local scope (voice-onset times, long and short vowel distinction, the marking of clauses, etc.). This view is supported by several aspects of language production when it is *spontaneous* (note that all studies have used written material). We mention early starts of sentences due to role-keeping found in conversational analyses ("So I — eh — where I started . . ."), repairs in mid-clausal position (Levelt, 1983) and the addition of sentence modifiers at the end of a clause that is syntactically complete ("I travelled — by train — uh — in France . . ."[2]). All these imperfections with which language production is abundant point to a property of the constituents of spontaneous speech: each is in principle unfinished and open to further modifications in an incremental, piecemeal fashion. On this account, prosodic structures, temporal or intonational, are traces of local constraints of all possible sources and not themselves global plans that have to be mapped upon the preplanned sentence constituents (see also Kempen and Hoenkamp, in press).

[2] C van Wijk Speaking, writing, and sentence form Ph D dissertation, University of Nijmegen (forthcoming)

The fact that we observed several effects that are frequently reported in phoneme monitoring studies indicates an adequate sensitivity of our experimental procedure. Still, the experiment may have been ineffective in disturbing a rhythm-based anticipation. One could argue that the anticipation is a result of prosodic cues immediately in advance of the accented syllables. The pretarget words, the word as it was originally recorded and its spliced-in counterpart of different length, were both spoken as immediate precursors of the same accented syllable. Therefore one could claim that by exchanging these words we also transferred sufficient prosodic cues as to when the accented syllable was to be expected, as was suggested by Dr C. Darwin (personal communication). However, we think that this line of reasoning is in conflict with the essential role of rhythm, namely a source of *global* redundancy. In fact, the anticipated time frame would then be so dependent upon segmental information in the one or two syllables in advance of the target that the essential advantage of such an extra source of timing would seem to be lost.

Reasoning along somewhat similar lines, Dr A. Cutler, in a review of this paper, supposed that rhythm may be confined to "clauses", such that temporal manipulation of a pretarget word located in one clause should have no effect on the processing of a target word in the next clause. In fact, in 15 of the 24 sentences a naive listener indicated a phrase boundary immediately before the target. If Cutler's suggestion is right, we ought to find a (small) effect from the temporal manipulation in the remaining 9 sentences (nos. 4, 8, 11, 15, 17, 18, 20, 22 and 23). However, despite the relative coherence of the prosodic environment of these targets, no RT disadvantage was observed for the manipulated conditions ($m1 = 623$ msec) compared to the unmanipulated conditions ($m2 = 628$ msec).

Clearly, much uncertainty seems to remain about the role of rhythm or temporal ordering in speech, both with respect to which events carry rhythm and with respect to the size of rhythmical structures: the sentence, clause or even group of words. Therefore, there is no consensus between researchers as to which experimental effects do or do not count as evidence for rhythm. It is in the light of this problem that we want to discuss a study of Buxton (1983) who manipulated the temporal structure of spoken sentences in a manner very similar to ours. In all sentences Buxton cross-spliced between short noun phrases such as "... with the/blue toy ..." and "... with the/blueish toy ..." on the one hand, and "... with the/red toy ..." and "... with the/reddish toy ..." on the other hand (bars indicating points of splicing). She found long phoneme-monitoring RTs when a one-syllable word was replaced by a two-syllable word (blue→reddish) or vice-versa (blueish→red) and short RTs when the exchanged words had an equal number of syllables (blue→red; blueish→reddish).[3] Of course Buxton took this result as supporting Martin's idea. Two aspects of Buxton's results, however, attract attention. First, the effect was equally present very early (e.g. after only two short words) as it was late in a sentence; second, no effect of word length was obtained (long pretarget words should elicit shorter RTs, and this was found by us, see Table II). To what extent could the fact that she spliced within such phrases have caused local disruptions instead of the global effects appropriate for testing the existence of rhythm? There are many dependencies between neighbouring syllables, even leading to the notion of "coproduction" (Fowler, 1981). "Coproduction" refers to a hypothetical superimposition of the production of unstressed vowels on a stressed vowel. It is intended to explain the shortening of the stressed vowel and the anticipatory and carryover coarticulatory effect found in the timbre of the unstressed vowels that surround it. The shortening of a stressed vowel is well established: the duration of a syllable decreases with an increasing number of other syllables that are part of the same word (Nooteboom and Doodeman, 1980) and even the same clause (Huggins, 1975). Tentatively, we argue that splicing within such phrases has the unwanted effect of damaging the tight articulatory/prosodic unity of these short phrases. Particularly, changing the number of syllables could result in a mismatch of durational features between the first and the second part of the coproduced entity. Problems in processing the pretarget word would then be the main cause of the RT delay to the targets. This local disruption can indeed be expected to overrule normal word-processing effects and to be equally strong early as it is late in a sentence. None the less, further research is needed because of the close similarity between Buxton's study and ours and the lack of explicit criteria of what qualifies as a rhythmic event and what as a manipulation of rhythm.

Finally, the possibility should be considered that the sentences investigated in this study simply lacked the rhythmic properties that the material of Meltzer et al. (1976)—among others—did possess. In other words, rhythm may be language specific, even to the extent that it is present in one supposedly stress-timed language (English) and not in another one (Dutch). We do not want to exclude this possibility. Such a proposition, however, again weakens the notion of rhythm, making it

[3] One might wonder whether we would have found an effect of temporal distortion if we had used pretarget words with the same syllabic structure as Buxton has. However, in 15 of the 24 sentences pretarget words *were* of the type "strong" and "strong–weak" as those of Buxton. The mean RTs of this subset were quite similar to the mean overall RTs. These 15 sentences yielded for Condition 1: 601 ms; 2: 572 ms; 3: 577 ms; 4: 605 ms.

less fundamental and universal to speech perception than suggested by allusions to an "analysis-by-synthesis" theory (Martin, 1972) or to it being an "integral part of the sentence comprehension process" and making "eminent strategic sense", even changing the left-to-right processing order (Cutler, 1976).

In sum, given the results reported in this paper and our tentative interpretation of the results from the other studies discussed, we are inclined to conclude that speech production on a global level is a process of concatenation. This view is in line with findings of Maassen and Povel (1984) and Osberger and Levitt (1979) showing that the artificial correction of the temporal structure of deaf speech has little positive effect on intelligibility. Moreover, it fits within the model of Povel (1981, 1984) and Povel and Essens (1985), which states that the temporal structure in a serial pattern is only conceptualized accurately if the pattern, on the basis of its accent distribution, evokes some regular internal time-base reflected in a tendency to tap a regular beat along with the pattern. Listening to speech seldom seems to elicit such a tendency.

REFERENCES

Allen, G. D. (1972). The location of rhythmic stress beats in English: An experimental study I. *Language & Speech*, **15**, 72–100.

Buxton, H. (1983). Temporal predictability in the perception of English speech. In A. Cutler and D. R. Ladd (Eds.), *Prosody: Models and measurements*. Berlin: Springer-Verlag.

Cutler, A. and Norris, D. (1979). Monitoring sentence comprehension. In W. Cooper and E. Walker (Eds.), *Sentence processing*. New York: Laurence Erlbaum Associates.

Cutler, A. (1976). Phoneme monitoring RTs as a function of preceding sentence contour. *Perception & Psychophysics*, **20**, 55–60.

Donovan, A. and Darwin, C. J. (1979). The perceived rhythm of stress. In E. Fischer-Jørgensen, J. Rischel and N. Thorsen (Eds.), *Ninth International Congress of Phonetic Sciences*, Copenhagen.

Fowler, C. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research*, **24**, 127–139.

Fowler, C. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General*, **112**, 386–412.

Haggard, M. (1971). *Speech synthesis and perception* (Progress Report No. 5). MRC Applied Psychology Laboratory, Cambridge, England.

Huggins, A. (1975). On isochrony and syntax. In G. Fant and M. Tatham (Eds.), *Auditory analysis and perception of speech*. London: Academic Press.

Kempen, G. and Hoenkamp, E. (in press). An incremental procedural grammar for sentence formulation. *Cognitive Science*.

Lehiste, I. (1973). Rhythmic units in production and perception. *Journal of the Acoustical Society of America*, **54**, 1128–1234.

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition*, **14**, 41–104.

Maassen, B. and Povel, D.-J. (1984). The effect of correcting temporal structure on the intelligibility of deaf speech. *Speech Communication*, **3**, 123–135.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, **79**, 487–509.

Meltzer, R., Martin, J. G., Bergfeld Mills, C., Imhoff, D. and Zohar, D. (1976). Reaction time to temporally-displaced phoneme targets in continuous speech. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 277–290.

Nooteboom, S. and Doodeman, G. (1980). Production and percption of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, **67**, 276–287.

Ohala, J. J. (1975). The temporal regulation of speech. In G. Fant and M. Tatham (Eds.), *Auditory analysis and perception of speech*. London: Academic Press.

Osberger, M. and Levitt, H. (1979). The effect of timing errors on the intelligibility of deaf children's speech. *Journal of the Acoustical Society of America*, **66**, 1316–1324.

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology. Human Perception and Performance*, **7**, 3–18.

Povel, D.-J. (1984). A theoretical framework for rhythm perception. *Psychological Research*, **45**, 315–337.

Povel, D.-J. and Essens, P. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411–440.

Uit Den Boogaart, P. (1975). *Woordfrequenties in geschreven en gesproken Nederlands*. Utrecht: Oosthoek, Scheltema & Hoeksema.

Warren, R. and Obusek, C. (1971). Speech perception and phonemic restoration. *Perception & Psychophysics*, **9**, 358–362.

Revised manuscript received 19 August 1985

APPENDIX

List of Experimental Sentences

Nouns as pretarget words, 16 sentences. Target phonemes are in upper case.

1. Jan heeft slordig gekookt want er dreef een stukje kurk/kn̅o̅flŏok Bovenop de soep
2. Het meesterlijk talent van de componist hoor je hier in een klank/ke̅n̅mĕrk Betoverend door zijn orginaliteit

3. Toen we hem in volle actie zagen viel de soepelheid van jullie kat/kăndĭdāāt Best wel op
4. Allereerst schoof hij een kast/kāpstŏk Binnen door de tuindeuren van ons nieuwe apartement
5. Om in de kelder toch wat ruimte over te houden kun je het wijnrek en die kist/kōēlkăst Beter aan de muur monteren
6. Kris heeft potverdorie met het verkeerde gereedschap de kap/kōplămp Botweg achterstevoren gemonteerd
7. 's Winters maakt Juultje van alle sneeuw in de tuin een pop/pūīnhŏop Buiten alle proporties
8. Ik vind persoonlijk voor het vraagstuk van militaire oefenterreinen die plek/pŏlĭtīēk Bepaald niet de goede keus
9. Op oma's nachtkastje ligt sinds zijn dood ome sjaaks pet/pŏrtrēt Binnen rijkwijdte voor als ze heimwee krijgt
10. Wanneer uw financiele toestand ons dwingt moeten we wel het tekort uit uw pot/prŏjēct Bij de amro zien te halen
11. Dat pas twee dagen later mijn post/pāspŏort Binnengleed vind ik wel tekenend voor de ptt
12. Die groningse popgroep lukt het toch weer om met een prima plaat/prŏdūctië Bovenin de hitlijst te staan
13. Het was van Utrecht naar Maastricht toch eigenlijk wel een tocht/trănspōrt Boven onze krachten
14. Ik zie heel wat reactionair geschrijf maar dat is een tekst/tījdschrĭft Bij het fascistische af
15. Als je eens nauwkeurig zijn taak/tākĕnpăkkēt Beoordeelt zie je dat hij eigenlijk een slimme vent moet zijn
16. Gelukkig werd hij ziek anders had hij toch nog mijn tip/tījdstĭp Bijna doorgegeven

Verbs as pretarget words, 8 sentences.

17. Sinds zijn promotie is 'ie uiterst verwaand geworden en hij pikt/prĕstēērt Bijna niets meer
18. Hoe dan ook de ministerpresident perst/părkēērt Binnenlandse-zaken in een onmogelijke positie
19. Vreemd dat hij het plan voor het stadhuis nu toch toont/tōēlĭcht Bij de raadsvergadering
20. Ma heeft gezegd dat als hij toch zo nodig tikt/tīmmĕrt Bovenop zolder dat hij dan het luik dicht moet doen
21. Ik vind het wel in orde dat een coach z'n team streng traint/tōēsprĕekt Binnen zekere fatsoensgrenzen
22. De verknipte taalkundige aait stotterende duiven en knijpt/kōēstĕrt Bilinguistische marmotten

23. Hongersnood vitaminetekort en vergaand longoedeem kwelt/ke̅n̅mĕrkt Biljoenen mensen elk jaar weer
24. 't Is zeker dat al wat Jonst kookt/kla̅a̅r̅măakt Buitengewoon lekker is en zo vol van aandacht

Epilogue

In the introduction three main issues were formulated: the inherence of organization, the completeness of structural representations and the autonomy of primary perception of structure. We will start our present concluding discussion by reviewing what the five studies presented in this thesis contribute to these issues. Then we will examine a possible reason for the lack of context effect reported in Part III, referring to recent discussions about the informational encapsulation of perceptual processes. Finally we will assess what these studies add to the value of Structural Information Theory as a model for human perception.

Contributions to the three main issues

Inherence. The perceived organization of the components within a (simple wire-like) pattern is not an aspect secondary to the perceptual representation, not derived in a post-access evaluation. In our studies using a fast presentation of a complete pattern and a subpattern, we eliminated post-access confoundings by looking at the interaction between order of presentation and type of subpattern (chapters 1 and 2) and by applying a bias correction (chapter 2). In this way, we demonstrated that it is likely that perceived organization is an integral aspect of the initial representation, or primary perception, of structure. SIT codes and SIT's selection criterion (the minimum principle) appeared to have predictive power as regards the early processing effects of chapters 1 and 2. Therefore, the first study could justifiable be called a study on "knowledge within perception".

Completeness. In chapter 2 it was demonstrated that a non-preferred "hidden" representation is indeed concurrently present in primary perception, rather than being the basis of just another view alternating with the preferred one. This is in accordance with the theoretical incompleteness of the preferred interpretation and the accessory role of a special "hidden" representation: the complementary code. The concurrent presence could be concluded from the large effect of presentation order on the bias-corrected recognition scores of subpatterns corresponding to that complementary representation. There is one minor discrepancy between the results presented in chapter 1 and those in chapter 2. In contrast to the first study, the effect of presentation order reported in the second study appeared already upon the near-simultaneous presentation of complete and subpattern, which invalidates the hypothesis formulated in chapter 1 that pattern interpretations would be developed at a rate of 30 ms. The presently used method may not allow a definite measurement of processing times. In chapter 3 a special type of complementary code was tested. Subjects had to discriminate between arrays that varied only with respect the number of (identical) elements, which is a metrical rather than a structural aspect. The complementary codes proposed in chapter 3 allow for a complete representation of the number of elements by means of the specification of the relative position of the elements. The complexity of these codes increases drastically when going from three to four elements. The reaction times of the subjects in a same/different task were in agreement with the code complexity. It was argued that "subitizing" (the seemingly immediate apprehension of the numerosity of a small number of elements compared to the difficulty encountered with larger numbers) can be understood in terms of these metrical codes.

Autonomy. The first three chapters show primary perception to be "smart": it seems to be capable of finding the single best or minimum code, while securing completeness of representation by means of a complementary code. In the last two studies, primary perception appeared to be isolated, or "autonomous", from information from sources other than the immediate stimulus. In other words, we only found an effect of "knowledge within perception" but not of "knowledge on perception". In chapter 4 prolonged processing, reasoning, beliefs etc. did have an effect on the recognition of subpatterns. However, when biases were properly controlled, no indications were found that expecting a particular subpattern altered the initial organization of the complete pattern. In a control condition, expectancy did appear to allow for selective attention to local features of the complete pattern. The conclusion was that context effects on primary perception have to be discerned from biases and selective attention. In chapter 5 we reported a similar autonomy of context for the domain of word recognition. The investigated "context" in this case is the global timing pattern of the sentence. According to Martin (1972), the global timing pattern of sentences in stress-timed languages such as English and Dutch should lead to an active expectancy and therefore to a speeded word recognition. Again, local aspects of the stimulus (frequency and length of the word) did affect the results in contrast to the wider context. One word of caution is required: we tested the alleged predictive role of rhythm, instead of the question whether or not our sentences contained a rhythmic timing in the first place. In fact, in chapter 5 we argued *against* the hypothesis that speakers would be able to impose a regular timing upon their normal, spontaneous productions. If speakers do not impose timing, it is of course incorrect to speak of an autonomy from context: there would be no context to start with.

Why autonomy?

With respect to our conclusion that the primary perception of structure proceeds in an autonomous fashion, it is important to note that meaningless wire-like patterns were used. For other types of patterns, special-purpose processing mechanisms may be effective (cf. Calis & Mens, 1985) which may or may not operate autonomous. However, in the absence of relevant data, we consider it plausible that the presently established autonomy will hold in general. Nevertheless, it needs to be stressed that context is important for the prolonged processing of complex displays. Such effects, however, may be called post-access compared to the primary perception of objects that can be perceived in a single glance.[1]

Concentrating on the studies on visual organization, two major conclusions can be made. On the one hand, chapters 1 and 2 show primary perception to be an extremely efficient mechanism which succeeds in producing the complete set of the best and the complementary representation of an entire

---

[1] Indeed, there are well-known examples of informed recognition, e.g., observed in R. James's picture of a well-camouflaged dog which remains a meaningless array of blobs until the idea is born that it might be a photograph of a dog. However, these are exceptions and require prolonged processing. Rather than proving that the border between "perception" and "cognition" is diffuse, these demonstrations may well be taken to demonstrate that in the absence of a stable primary shape organization, only scrutiny and selective attention can replace the normal effortless processing.

pattern, possibly within a temporal resolution as small as 30 ms. On the other hand, in chapter 4 primary perception shows its limitations by not taking into account relevant information from a stimulus shown just a few seconds earlier. Why should such an efficient system be so "short-sighted"? Is autonomy an accident of nature, an advantageous property, or a necessity for any object-recognition to occur?

Fodor (1983) proposes that perceptual processes can be fast and efficient *because* they are not burdened by every piece of knowledge that might be relevant to the situation at hand: perceptual processes are "informationally encapsulated". Thus, context-insensitivity in general is considered advantageous, on this account. Indeed, everyone would agree that, for instance, shape perception cannot be based on an evaluation of all beliefs and yet be as fast as it is. Moreover, one may hold that normally, outside a priming task, one's beliefs are quite unreliable and seldom predictive (Fodor, 1985, p. 5). For example, one may expect to see a table but chances are that it is something else, for instance a chair. However, we found more than an autonomy with respect to relatively high-level knowledge of the world. The question now is: why should it be efficient for the process resulting in a pattern organization to be so completely isolated from prior information, even in case of the extreme compatibility of prior and new information that has been described in chapter 4?

Apart from the fact that one may mistakingly expect a particular object and/or event, there may be another reason why prior information may seldom be predictive. The number of ways that an object can present itself visually is enormous. For many objects, it even seems to be impossible to give one general structural description. The set of chairs, again for example, seems to be expandable infinitely with structurally different exemplars.[2] But even if prior information is structurally compatible with the input, there is still the simple fact that the visual information may vary dramatically on a lower level, where it is altered by changes in size of the object, orientation, illumination, interposition, type of contour, etc. In the Introduction we noted that object recognition is remarkably indifferent with respect to the lower-level

---

[2] Biederman (1987) argues that the number of all "readily discernible" exemplars is limited. However, it may be questioned whether a "geon", his basic unit of which objects such as a chair are assumed to be composed, cover structurally identical patterns; considerable structural variation is possible within each geon-class.

variations. For example, a contour in drawings is generally reduced to a line (that often corresponds to a physically impossible contour; Koenderink & Van Doorn, 1982); in comic strips detailed surface information (e.g., of the texture of the skin) is often replaced by a uniform colour. Indeed, object recognition is often quicker from schematic drawings than from photographs (Ryan & Schwartz, 1956). This may be interpreted to show that structure is recovered in a flexible way along many alternative routes, none of which is essential to recognition. Accordingly, students of computational vision nowadays express the idea that an effective vision system probably employs many sources of visual information (quantitative and qualitative), as opposed to Marr who focussed on quantitative surface information. Witkin and Tenenbaum (1983), for example, stress the importance of structural properties, such as the direction of shading and texture gradients relative to boundary contours, junctions, highlights, parallelism, symmetry and repetition. Given many alternative clues to shape, the effectiveness of background knowledge, and context in general, would seem questionable. In other words, an independent level of classification on the basis of structure may often be a *necessity, in precedence* to any classification in terms of meaningful objects, at least in situations that are not (artificially) restricted.

It may be instructive to make the same point about the relative use of context with respect to auditory word-recognition. Shape organization may be compared to the identification of phonemes (or syllables, or whatever closed-class sub-word recognition-unit one favours), as has been done by Biederman (1987). What we now want to argue is that word recognition can only benefit from context (for instance, by matching highly activated entries first) *above* the phonemic level; below that level, there may be too many combinations to consider when linking a (semantic) expectation to the acoustical signal.

Fodor (1983) takes a somewhat different position by suggesting that informationally encapsulated systems are an accident of nature yielding an advantage in surviving: such systems allow fast responses ("a tiger! run away!") without confusion due to beliefs that may be wrong (such as: "tigers don't eat psychologists"). Indeed, one may note that there are two ways in which the visual system can be "flexible". Adapting to task demands and prior knowledge of our world is one type of flexibility. Sensitivity to and speed of processing of the (unexpected) immediate situation (whatever "world" it may be!) is another one. The latter type of flexibility clearly may not suffer from the first type.

SIT and models of perception

It is a basic research strategy of SIT to predict the perceptual interpretation of visual and auditory patterns by means of a specification of stimulus structure. In other words, the basic assumption is that an interesting part of choice behaviour in visual and auditory tasks can be understood from constraints imposed by SIT on the perceptual representation of structure. Thus, SIT as such is a theory of representation and not one about specific processing mechanisms which lead to the perceptual interpretation. In spite of the success of this approach in many studies, the lack of a plausible process model has often been brought forward as a reason why SIT would not do as a complete theory of perception (Kolers, 1983; Chase, 1986; Pomerantz & Kubovy, 1986; Hatfield & Epstein, 1985). One might be tempted to think that the presently reported relevance of SIT to predict early processing effects may be an answer to this critique. However, this is not what we want to claim.

Let us first try to formulate what makes a model a *process model.* An acceptable description may be that a process model is a model of which it is claimed that all of its essential entities (operations, stages) are based either on what currently is generally accepted knowledge of the psychological mechanism, or are at least directly falsifiable in a psychological experiment. It may be clear that we have not tested all the essential operations and/or stages of SIT. In particular, we still do not know *how* the visual system arrives at the representations that we have probed in our experiments. The encoding scheme described in the Introduction still is not a process model. [3] In the past, several widely varying types of processes have been proposed to complement SIT: Restle (1982) considered a distributed relaxation process. Buffart and Leeuwenberg (1982) compared codes with hypotheses and envisaged

[3] Part of the encoding scheme has recently been implemented in a computer algorithm that always finds the *minimum code* for each *primitive code* (Van der Helm & Leeuwenberg, 1986). What this shows with respect to the human visual system, is that such a reduction is possible. However, it does not show *how* the visual system performs it (or even whether it performs something comparable in the first place).

a cyclic testing of increasingly better codes. Van Leeuwen, Buffart and Van der Vegt (1988) built a network memory on the generic relations between codes supplemented by a few assumptions about the spreading of activation within the network; again patterns are recognized in a "top-down" testing procedure. The latter approach has already been successful in simulating effects of viewing span and practice on the preference for other codes than the best code found in experiments with complex sequential patterns (see also Van Leeuwen & Buffart, 1987). In short, SIT does provide a fruitful context to formulate process models. It is interesting that developments in this area do not exclusively build on the application of reduction rules and a sequential primitive code. Instead, inspiration comes from the general principles of SIT. Some of these principles were described in the Introduction: the maximal specification of regular identity relations and the incompleteness of the resulting representations. Some principles are still in statu nascendi: the stability of network nodes (Van Leeuwen, Buffart & van der Vegt, 1988), hierarchical completeness (Buffart, 1987) and accessibility (Van der Helm & Leeuwenberg, 1988). And, of course, future investigations may yield other, perhaps even more fundamental insights.

What the present results contribute to this issue is the following. Coding theories in general are developed to function as normative models of choice behaviour pooled over many trials and subjects under different states of learning, employing different (unknown) strategies. Indeed, an individual response in an experiment is subjected to extraneous factors, such as fatigue, or ideosyncratic associations, which from time to time may prevent the "best" code to be reached. Because of such factors, it has even been argued that SIT predicts "late perception", that is, it would not apply to "early perception" (Buffart, 1987; see also Restle, 1982, p. 33). Buffart described "early perception" as the first 5 seconds after stimulation needed to reach a stable final interpretation. Yet, both in chapter 1 and 2 specific predictions based on SIT were confirmed by early processing effects that could be ascribed to primary perception. This does not imply that SIT is a process model, but SIT does seem to be valid for understanding the format of early perceptual representations, at least within the constraints presently obeyed: using simple meaningless patterns viewed in a single glance. The representation of structure described by SIT, rather than just stemming from the realm of reasoning and fantasy, appears to have a bearing on primary perception.

References

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review, 94*, 115-147.

Buffart, H. & Leeuwenberg, E. (1982). Structural Information Theory. In: H.-G. Geissler, H. Buffart, E. Leeuwenberg & V. Sarris (Eds.), *Current issues in perception*. Amsterdam: North-Holland Publishing Company.

Buffart, H., Leeuwenberg, E., & Restle, F. (1983). Analysis of ambiguity in visual pattern completion. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 980-1000.

Buffart, H. (1987). Seven minus two and structural operations. In: E. E. Roskam, & R. Suck (Eds.), *Progress in Mathematical Psychology, I*. Amsterdam: North-Holland Publishing Company.

Buffart, H. (1987) In defense of overall figural goodness. Internal Report 87 NICI 01, University of Nijmegen.

Calis, G., & Mens, L. (1985). Primary stages of single-glance face recognition: Expression and identity. In: H. Ellis, M. Jeeves, F. Newcomb & A. Young (Eds.), *Aspects of face processing*. Dordrecht, N.L.: Martinus Nijhoff.

Chase, W. G. (1986). Information processing. In: K. Boff, L. Kaufman & J. Thomas (Eds.), *Handbook of perception and human performance, Vol. II: Cognitive processes and performance*. New York, NY: John Wiley and sons.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, M.: The M.I.T. Press.

Fodor, J. A. (1985). Precis of The modularity of mind. *The behavioural and brain sciences, 8*, 1-5.

Hatfield, G., & Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin, 2*, 155-186.

Koenderink, J. J., & Van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception, 11*, 129-137.

Kolers, P. (1983). Perception and representation. *Annual Review of Psychology, 34*, 129-166.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behaviour. *Psychological Review, 79*, 487-509.

Pomerantz, J., & Kubovy, M. (1986). Perceptual organization and cognition. In: K. Boff, L. Kaufman & J. Thomas (Eds.), *Handbook of perception and performance, Vol. II: Cognitive processes and performance*. New York, NY: John Wiley and sons.

Restle, F. (1982). Coding theory as an integration of gestalt psychology and information processing theory. In: J. Beck (Ed.), *Organization and representation in perception*. Hillsdale, N.J.: Lawrence Erlbaum.

Ryan, T. A., & Schwartz, C. (1956). Speed of perception as a function of mode of presentation. *American Journal of Psychology, 69*, 60-69

Seidenberg, M. (1985). Lexicon as module. *The [illegible]iou[illegible] and brain sciences, 8,* 31-32.

Sternberg, R. J. (1985). Controlled versus automa[illegible] pro[illegible]ing. *The behavioural and brain sciences, 8,* 32-33.

Van der Helm, P. & Leeuwenberg, E. (1986). Avoiding explosive search in automatic selection of simplest pattern codes. *Pattern Recognition, 19,* 181-191.

Van der Helm, P., & Leeuwenberg, E. (1988). Accessibility, a criterion for the choice of visual coding rules. Internal Report 88 NICI 01, University of Nijmegen.

Van Leeuwen, C., Buffart, H., & Van der Vegt, J. (1988). Sequential influence on interpretation of meaningless serial patterns: Local precedence and global preference. *Journal of Experimental Psychology: Human Perception and Performance,* in press.

Van Leeuwen, C. & Buffart, H. (1987). Facilitation of retrieval by cue structure; crossing the gap between perception and memory. Internal Report 87 FU 05, University of Nijmegen.

Summary

Theoretical accounts of stimulus structure can be used to predict the perceptual organization of stimuli. This thesis examines the perceived organization of simple line drawings using Leeuwenberg's (1971) Structural Information Theory (SIT) as a starting point. With respect to the recognition of spoken words, the importance of the rhythmic organization of sentences as predicted by Martin (1972) is investigated.

Three issues are considered. The first question, addressed in Part I, is whether the organization is **inherent** in perceiving a pattern, or instead, added to the initial registration, for instance by reasoning. To discriminate between these two possibilities in our experiments, effects of *primary perception of structure* are distinguished from secondary effects due to prolonged processing. Primary perception of structure is defined in this thesis as the initial access of the representation of structure; *post-access* effects are the result of subsequent semantic, associative or other memory processes. Referring to "primary perception of structure" does not mean that there are a-priori reasons to believe that the representation of structure is the result of a distinct process, separate from the rest of the system. Instead, the distinction between *primary* and *secondary* perception will turn out to be meaningful only for empirical reasons.

In Chapter 1, a task is introduced to eliminate such post-access confoundings. As in the embedded figures test, a complete pattern is presented followed by one out of several possible subpatterns. Recognition of the subpattern is measured at about threshold level: the complete pattern and the subpattern are each presented for only 10 milliseconds and they are separated by a very short interval. In chapter 1, the pairs are also presented in the reversed order. Pairs of a complete pattern and a subpattern are used which have a mutually exclusive organization when presented in one order. However, presented in the other order, the preferred representation of the first pattern—according to SIT—is able to reinforce a latent interpretation of the second one. By systematically varying the interval, a shortest interval of 30 ms is found at which the comparison between the two orders of presentation

results in a difference in the percentage correctly recognized subpatterns. It is concluded that this interval is related to the time to process the stimuli from a sensory registration into an organized representation of stimulus structure.

Considering this support for the inherence of organization in primary perception, a second and a third issue becomes relevant. The second issue, the focus of Part II, is about organization and the **completeness** of representations. Each pattern allows for alternative organizations that often cover different structural aspects of that pattern. This agrees with the theoretical incompleteness of a single code according to SIT. SIT assumes for most patterns a second code (apart from the best or *minimum code* which corresponds to the preferred interpretation). This second code, often the second-best, is called *complementary* because it covers all the structural regularities missing in the minimum code. In many earlier studies based on SIT the frequency with which the preferred and the second-best organization is reported (their respective "perceptual strength" or *Prægnanz*) could be related to the complexity of the minimum code relative to the complexity of the second-best code. The question is whether these two codes indeed relate to actual perceptual representations that are concurrently present.

In chapter 2 it is shown that the complementary second-best representation is indeed concurrently present in primary perception instead of being the basis of another view of the pattern alternating with the preferred one. As in Chapter 1, early processing effects are investigated by presenting a complete pattern and a subpattern in a very fast sequence. These subpatterns correspond to the best, the second-best or an odd organization of the complete pattern. The recognition scores are corrected in order to discriminate post-access effects (due to prolonged processing, demand-characteristics and other biases, for example in favour of the "better" subpatterns) from effects in the primary representation of structure. The correction is based on the distribution of the erroneous recognitions of subpatterns. The concurrent presence is evident in the corrected scores. A minor discrepancy is that the straightforward conclusion of chapter 1 about the rate of 30 ms at which pattern interpretations would be developed is not confirmed by this study. Therefore, varying the interval in such a task does not seem to yield a reliable indication of central processing times.

A structural code may not only be incomplete with respect to the full structure of a pattern, it does not—by definition—cover metrical aspects. The task described in chapter 3 requires an accurate perceptual representation of the *number* of identical elements: a more or less metrical aspect. Therefore, a special type of complementary code is considered. Tentatively, the relative position of pairs of elements is completely specified with binary ordering relations. Higher-level orderings can be made between pairs that form a perceptual group. Subjects in the experiments have to indicate as fast as possible whether two successively presented arrays contain an equal number of elements. The reaction times are in agreement with the abrupt increase in complexity of the corresponding codes for arrays of three elements versus four elements. It is argued that "subitizing" (the seemingly immediate apprehension of the numerosity of a small number of elements compared to the difficulty encountered with larger numbers) can be understood in terms of the present metrical codes.

The third issue is addressed in Part III, consisting of chapters 4 and 5. In the preceding chapters, the visual system was shown to represent the internal structure completely. In the last two chapters it is investigated whether the "external structure" of a pattern—in the present case the wider context of earlier perceived patterns—is equally important. In particular, our third question is whether primary perception can depend on context. This is called the issue of the **autonomy of organization.** In chapter 4, many experimental manipulations are combined in an attempt to bias primary perception in favour of one or another organization of a simple line pattern. Subjects are led to expect a particular organization by starting a trial with a special *prime*: a pattern in a segmented (exploded view) version corresponding to but one of the organizations of the complete pattern shown a few seconds later. The question is whether this prime, and the expectation it should evoke, changes the initial organization of the complete pattern. Again the perceptual representation of the complete pattern is probed by looking at the recognition of subpatterns presented after a very short interval (30 ms). It is concluded that prolonged processing, reasoning, beliefs etc. may in part determine what organization a subject will prefer eventually, but that if biases are properly controlled, no indications are present that an expectation of a particular subpattern alters the initial organization of the complete pattern. It is argued that context effects on perceived organization, reported in the literature, are either

due to a post-access evaluation of stimuli or to a task simplification leading to selective attention for a local feature.

In chapter 5 a similar autonomy is reported for the recognition of spoken words in Dutch sentences. Dutch speakers as well as speakers of other so-called "stress-timed" languages, appear to impose a certain rhythmical structure on their sentences. Martin (1972) claimed that such a rhythm provides a reliable context to the listener. Once locked into the rhythm of a sentence, a listener would be able to anticipate the important words to follow. In the present study, subjects had to react as fast as possible upon the occurrence of a specific phoneme. Such a *phoneme monitoring* task is relatively free of post-access confoundings. These phonemes were placed at the beginning of stressed words. The global temporal pattern was manipulated by replacing the word just in advance of the word with the target phoneme by another one differing in length. Local stimulus aspects (word length and level of stress) did affect reaction times; this in contrast to the manipulations of the temporal context which did not produce any effect upon the detection of the phoneme. The conclusion is that the recognition of spoken words is not guided by a global temporal organization of the embedding sentence.

In the Epilogue two topics are discussed. In the first place, an understanding of the autonomy established in chapter 4 is sought in the possibility that contextual information is of little use for object recognition. An intermediate level of classification (*structure*) may be advantageous and even necessary *in precedence* to any classification in terms of meaningful objects. In the second place, it is argued that the present results do contribute to the plausibility of SIT as a theory of perception. SIT, being primarily a theory of perceptual representations, has recently been criticized for the lack of a particular notion by what processes those representations are reached. This thesis shows that SIT is valid for understanding the format of early representations; SIT does appear to bear on *primary perception*.

Samenvatting

Theoretische beschrijvingen van de structuur van stimuli kunnen gebruikt worden om de waargenomen organisatie van die stimuli te voorspellen. In dit proefschrift wordt de waargenomen organisatie van simpele lijnfiguren onderzocht met voornamelijk Leeuwenberg's Structurele Informatie Theorie (1971) als uitgangspunt. Ten aanzien van de herkenning van gesproken woorden wordt het belang onderzocht van de ritmische organisatie van zinnen zoals voorspeld door Martin (1972).

Drie kwesties worden beschouwd. De eerste vraag, die wordt behandeld in Deel I, luidt: is organisatie **inherent** aan het waarnemen van een patroon of slechts toegevoegd aan de aanvankelijke perceptuele representatie, bijvoorbeeld door middel van redeneren? Om in experimenten hierop antwoord te kunnen geven, worden effecten van de *primaire perceptie van structuur* onderscheiden van secundaire effecten ten gevolge van verdergaande verwerking. Primaire waarneming van structuur wordt in dit proefschrift gedefinieerd als de eerste toegang (*access*) tot de structuurrepresentatie van de stimulus; *post-access* effecten zijn het resultaat van daaropvolgende geheugenprocessen van semantische, associatieve of andere aard. Het gebruik van de uitdrukking "primaire perceptie van structuur" betekent overigens niet dat er a-priori redenen zijn om aan te nemen dat de representatie van structuur het resultaat is van een duidelijk onderscheiden proces, apart van de rest van het systeem. Integendeel, het onderscheid tussen *primaire* en *secundaire* waarneming zal zijn betekenis proefondervindelijk vinden.

In hoofdstuk 1 wordt een taak geïntroduceerd welke post-access effecten dient uit te sluiten. Zoals in de *embedded figures* test wordt een complete figuur aangeboden en daarna één van enkele mogelijke deelfiguren. De herkenning van de deelfiguur wordt gemeten op drempelniveau: de complete figuur en de deelfiguur worden elk slechts 10 milliseconden aangeboden met daartussen een zeer kort interval. In hoofdstuk 1 worden deze paren figuren ook in de omgekeerde volgorde aangeboden. De paren van figuren die gebruikt worden hebben onverenigbare organisaties in de ene aanbiedingsvolgorde. Echter, omgekeerd aangeboden kan bij deze paren—volgens SIT—de geprefereerde of *minimum* representatie van de eerste figuur een verborgen interpretatie van de tweede figuur aanspreken. Door het interval systematisch te varieren is 30 milliseconden als kortste tijd gevonden waarbij de vergelijking van de twee aanbiedingsvolgorden een verschil opleverde in het percentage correct herkende deelfiguren. De conclusie is dat dit interval gerelateerd is aan de tijd die het kost om de stimuli van een eerste ruwe opslag te verwerken tot een georganiseerde representatie van de stimulus structuur.

Gezien deze ondersteuning van de hypothese dat organisatie inherent is aan primaire perceptie, wordt een tweede en een derde kwestie van belang. De tweede kwestie, die aan de orde komt in Deel II, is die van figuurorganisatie en de **volledigheid** van representaties. Elke figuur staat meerdere organisaties toe die ieder met andere structurele aspecten van die figuur overeenstemmen. Dit sluit aan bij de onvolledigheid van een enkele code volgens SIT. SIT veronderstelt dat voor de meeste figuren een tweede representatie of code voorhanden is, naast de beste of minimum code die overeenstemt met de geprefereerde interpretatie. Deze tweede code (vaak de op-een-na beste) wordt *complementair* genoemd omdat deze alle structurele regelmaat beschrijft die in de minimum code ontbreekt. In vele eerdere onderzoekingen kon het aantal malen waarmee de geprefereerde organisatie werd gerapporteerd (en daarmee haar perceptuele sterkte of *Prægnantie*) op basis van SIT voorspeld worden aan de hand van de relatieve complexiteit van de minimum code ten opzichte van die van de op-een-na beste code. De vraag is of deze twee codes daadwerkelijk overeenkomen met perceptuele representaties die gelijktijdig aanwezig zijn.

In hoofdstuk 2 wordt aangetoond dat de complementaire op-een-na beste code inderdaad overeenstemt met een representatie die gelijktijdig maar "verborgen" aanwezig is in de primaire perceptie, in plaats van dat het een andere interpretatie is die afwisselt met de geprefereerde interpretatie. Evenals in hoofdstuk 1 worden vroege verwerkingsstappen onderzocht door een complete figuur en een deelfiguur in een snelle opeenvolging aan te bieden. De deelfiguren passen of in de geprefereerde, of in de op-een-na beste, of in een vreemde organisatie van de complete figuur. De herkenningsscores werden gecorrigeerd om post-access effecten (ten gevolge van verdere verwerking, wellicht onbedoeld gestuurd door aspecten van de taak of onder invloed van andere voorkeuren, bijvoorbeeld voor de "betere" deelfiguren) te scheiden van effecten die ontstaan in de primaire representatie van structuur. Deze correctie is gebaseerd op de distributie van de foutieve herkenningen van deelfiguren. De gelijktijdige aanwezigheid blijkt uit de gecorrigeerde scores. De conclusie van hoofdstuk 1 aangaande het tempo van één per 30 milliseconden waarmee figuurinterpretaties ontwikkeld zouden worden, werd helaas niet bevestigd in dit experiment. Het variëren van het aanbiedingsinterval blijkt daarom geen betrouwbare aanwijzing te geven voor de verwerkingstijd.

Een structurele code beschrijft niet alleen meestal de structuur van een figuur onvolledig, maar verontachtzaamt—per definitie—metrische aspecten. De taak die in hoofdstuk 3 geïntroduceerd wordt vereist een nauwkeurige perceptuele representatie van het aantal van een reeks gelijkvormige elementen: een meer metrisch aspect. Hiervoor wordt een speciaal type van complementaire code voorgesteld: de relatieve positie van paren van elementen wordt volledig vastgelegd in binaire ordeningsrelaties; ordeningen op een hoger niveau zijn toegestaan tussen paren die een perceptuele groep vormen. In het experiment moesten de proefpersonen zo snel mogelijk aangeven of twee opeenvolgend aangeboden reeksen een gelijk aantal elementen bevatten. De reactietijden stemmen overeen met de abrupte toename in complexiteit van de codes van reeksen met drie elementen naar reeksen met vier elementen. Een stelling is dat "subitizing" (het schijnbaar onmiddellijk bevatten van het aantal van een kleine verzameling elementen vergeleken met de moeite bij grotere verzamelingen) begrepen kan worden in termen van de hier voorgestelde codering.

De derde kwestie wordt behandeld in Deel III, bestaande uit hoofdstukken 4 en 5. In Deel II bleek het visuele systeem de "interne" structuur van een figuur volledig te representeren. In de laatste twee hoofdstukken wordt onderzocht of de "externe" structuur—in dit geval de context die gevormd wordt door eerder waargenomen figuren—evenzeer bepaalt wat de primaire perceptie is. Dit is de vraag naar de **autonomie** van de primaire perceptie van structuur. In het onderzoek dat wordt beschreven in hoofdstuk 4 worden vele experimentele manipulaties gecombineerd teneinde primaire perceptie te sturen in de richting van bepaalde organisaties van een eenvoudige lijn-figuur. De proefpersonen worden ertoe aangezet een bepaalde organisatie te verwachten door hun eerst een speciale *prime* aan te bieden: een complete figuur opgedeeld volgens één van de organisaties van de kort aangeboden complete figuur die enkele seconden later volgt. De vraag is of deze prime, en de verwachting die opgewekt wordt, de aanvankelijke organisatie van het complete patroon bepaalt. Weer wordt de perceptuele representatie van de complete figuur bepaald door te kijken naar de herkenning van deelfiguren die na een kort interval (30 millisecoonden) aangeboden worden. De conclusie van dit onderzoek is dat verdergaande verwerking (redeneren, overtuigingen etc.) weliswaar mede bepaalt welke organisatie van de complete figuur uiteindelijk geprefeerd wordt, maar dat er na aftrek van antwoordvoorkeuren geen aanwijzingen zijn dat de verwachting van een bepaalde deelfiguur de aanvankelijke organisatie beinvloedt. Vandaar dat gesteld wordt dat context-effecten die in de literatuur beschreven worden het gevolg zijn ofwel van een post-access evaluatie, ofwel van een vereenvoudiging van de taak welke aanleiding geeft tot selectieve aandacht voor (lokale) kenmerken van de stimulus.

In hoofdstuk 5 wordt een vergelijkbare autonomie gerapporteerd van de herkenning van gesproken woorden in Nederlandse zinnen. Nederlandse sprekers, evenals sprekers van andere "stress-timed" talen, lijken hun zinnen een bepaalde ritmische structuur mee te geven. Martin (1972) beweerde dat zulk een ritme een betrouwbare context verschaft aan de luisteraar. Eenmaal in het ritme van een zin gevangen zou de luisteraar in staat zijn op belangrijke fragmenten te anticiperen. In de huidige studie moeten de proefpersonen zo snel mogelijk reageren wanneer zij een bepaalde spraakklank (een bepaald foneem) hoorden. Een dergelijke *phoneme-monitoring* taak is verhoudingsgewijs vrij van post-access neveneffecten. Het foneem vormt altijd het begin van een beklemtoond woord. De globale temporele context is gemanipuleerd door het woord vlak vóór het woord met het foneem te vervangen door een ander woord met datzelfde foneem maar met een andere lengte. Locale stimulus aspecten (woordlengte en mate van beaccentuering) blijken de de reactietijden wel te beînvloeden, de manipulaties van de temporele context hebben geen effect. De conclusie is dat de herkenning van gesproken woorden niet gestuurd wordt door een globale temporele organisatie van de omringende zin.

In de Epiloog worden twee punten besproken. Op de eerste plaats wordt gepoogd autonomie te begrijpen door te overwegen dat mogelijkerwijs informatie vanuit de context van weinig waarde is voor objectherkenning.

Een tussenliggend niveau van classificatie (te weten dat van structuur) zou voordelig kunnen zijn en wellicht zelfs *noodzakelijk* om tot classificatie in termen van betekenisvolle objecten te komen. Op de tweede plaats wordt beargumenteerd dat de resultaten gerapporteerd in dit proefschrift bijdragen aan de aanvaarbaarheid van SIT als een theorie over de waarneming. SIT is in eerste instantie een theorie over perceptuele representaties en is recentelijk aangesproken op het ontbreken van een aanduiding volgens welke processen deze representaties bereikt worden in het waarnemingssysteem. Dit proefschrift verschaft zulk een aanduiding evenmin, maar het laat wel zien dat SIT relevant is om de aard van representaties vroeg in het waarnemingsproces te begrijpen: SIT blijkt betrekking te hebben op *primaire perceptie*.

Curriculum vitae

Lucas Mens werd geboren te Vught op 20 december 1956. Na het eindexamen Gymnasium β aan gymnasium Beekvliet te Sint-Michielsgestel studeerde hij vanaf 1976 Psychologie te Nijmegen, waar hij in 1983 het doctoraal examen aflegde met als hoofdrichting Functieleer. Hieraan voorafgaande assisteerde hij bij onderzoek naar spraakwaarneming en naar afasie op het Max-Planck Instituut te Nijmegen. Hij studeerde af op een studie naar de waarneming van ritme in spraak waarvan in hoofdstuk 5 verslag wordt gedaan, en op een eerste variant van het werk wat in dit proefschrift in hoofdstuk 1 beschreven is. Vervolgens was hij tot eind 1987 medewerker bij het ZWO project Structurele Informatie Theorie. Vanaf 1 october 1988 is hij als experimenteel psycholoog verbonden aan het binnenoorprothesen project van de afdeling KNO van het Sint-Radboud ziekenhuis te Nijmegen.

Stellingen

1. Van de in dit proefschrift aangetoonde aanwezigheid van complementaire interpretaties zijn we ons nauwelijks bewust.

2. Voor de classificatie en het onthouden van een patroon zijn de structurele Minimum code en Complementaire code voldoende; echter, voor het verkrijgen van een levendige voorstelling van dat patroon dienen deze codes omgezet en metrisch verrijkt te worden tot één enkele representatie.

3. Kennis eerder verworven dan 30 milliseconden voor aanbieding van een patroon heeft geen invloed op de initiële perceptuele organisatie, maar wel op een secundaire evaluatie van dat patroon. Deze twee momenten zijn te onderscheiden door middel van een response-bias correctie. *Dit proefschrift.*

4. Dat het verstaan van spraak niet afhangt van ritme is in overeenstemming met het feit dat een spreker een zin kan beginnen zonder dat het einde ervan voorbereid is. *Dit proefschrift.*

5. Identiteit en expressie van een gezicht zijn scheidbare perceptuele kenmerken. Het is namelijk met behulp van een zeer korte aanbieding van twee portretten mogelijk de lach van de een aan de identiteit van de ander te koppelen. *Zie G. Calis & L. Mens, 1985.*

6. Bij vele theoriën op het gebied van de visuele vormwaarneming waarbij het geldigheidsdomein is aangegeven dringt zich toch de vraag op hoe het visuele systeem voor een gegeven patroon kan bepalen of het tot het domein van die theorie behoort.

7. Het is te vrezen dat een verondersteld begrip van de menselijke geest in termen van een connectionistisch model zelf in fragmenten over deskundigen gedistribueerd zal zijn.

8. Bij de vormgeving van gehoortoestellen is het goed die van modieuze draagbare geluidsapparatuur te volgen. Aldus vermindert het acoustisch isolement van de walkmandrager het sociaal isolement van de gehoortoesteldrager.

9. Met deze stelling bent u het zeker niet eens.

Stellingen behorende bij het proefschrift:
Lucas Mens, *Primary perception of stimulus structure.* Nijmegen, 31 october 1988.